Missouri
DEPARTMENT OF ELEMENTARY & SECONDARY
EDUCATION™

End-of-Course Assessments
Technical Report
2018–2019

English II
Algebra I
Biology
English I
Algebra II
Geometry
Physical Science
Personal Finance

Submitted to the
Missouri Department of Elementary and Secondary Education

Presented by
Questar Assessment Inc.

Questar®

The Department of Elementary and Secondary Education does not discriminate on the basis of race, color, religion, gender, national origin, age, or disability in its programs and activities. Inquiries related to Department programs and to the location of services, activities, and facilities that are accessible by persons with disabilities may be directed to the Jefferson State Office Building, Office of the General Counsel, Coordinator – Civil Rights Compliance (Title VI/Title IX/504/ADA/Age Act), 6th Floor, 205 Jefferson Street, P.O. Box 480, Jefferson City, MO 65102-0480; telephone number 573-526-4757 or TTY 800-735-2966; fax number 573-522-4883; email civilrights@dese.mo.gov.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Below is a list of abbreviations that appear in this technical report.

ALD ....achievement level descriptor
ARC ....Assessment Resource Center
AYP ....adequate yearly progress
CLE...................course-level expectation
CR ....constructed response
CSEM ....conditional standard error of measurement
CTT...................classical test theory
DESE ....Department of Elementary and Secondary Education
DIF....................differential item functioning
DOK..................Depth of Knowledge
EFT ....embedded field test
ELL ....English language learner
EOC ....end-of-course
ESEA ....Elementary and Secondary Education Act
FRL ....free and reduced lunch
GLE...................grade-level expectation
GRF...................general research file
IAP....................Individualized Accommodation Program
IDEA.................Individuals with Disabilities Education Act
IEP ....Individualized Education Program
IRT....................item response theory
ISR ....Individual Student Report
ITS ....Internet Testing Systems
LEP ....limited English proficient
LOSS.................lowest obtainable scale score
MAP..................Missouri Assessment Program
MCDS ....Missouri Comprehensive Data System
MH....................Mantel-Haenszel
MOSIS ....Missouri Student Information System
NCLB................No Child Left Behind Act
PE......................performance event
RS ....raw score
RSS ....raw-to-scale score
SD ....standard deviation
SE......................standard error
SEM ....standard error of measurement
SR ....selected response
TAC ....Technical Advisory Committee
TE......................technology enhanced
WP ....writing prompt

Chapter 1: Introduction

1.1. Overview of Missouri End-of-Course (MO EOC) Assessments

The MO EOC assessments are standards-based assessments designed to measure students' knowledge of the Missouri Learning Standards, which define the knowledge and skills students need in each grade level and course for success in college, other postsecondary training, and careers.

1.1.1. MO EOC Content Areas

The MO EOC assessments include the following content areas:

- English I
- English II
- Algebra I
- Algebra II
- Geometry
- Biology
- Physical Science
- Government
- American History
- Personal Finance

In addition to the MO EOC assessments, the current Missouri Assessment Program (MAP) system includes the following assessment components for elementary and middle school:

- Grades 3–8 Communication Arts
- Grades 3–8 Mathematics
- Grades 5 and 8 Science

The statewide assessment program also includes the Dynamic Learning Maps (DLM) assessments in Communication Arts, Mathematics, and Science for students with severe cognitive disabilities, WIDA ACCESS for English language learners (ELLs), and a Personal Finance assessment for high school students who do not enroll in a personal finance course or who are receiving personal finance credit for embedded coursework.

1.1.2. A Brief History of MO EOC Assessments

English II, Algebra I, and Biology were developed and first administered in 2008–2009. English I, Algebra II, Geometry, Government, and American History were developed and first administered in 2009–2010. Physical Science was first administered in 2014–2015.

Table 1.1 provides the major events that have occurred for the MO EOC assessments from 2008–2009 to 2018–2019 to assist with the understanding and interpretation of test results throughout this report.

Table 1.1. Summary of Major Events from 2008–2009 to 2018–2019

Accountability Year	Event(s)
2008–2009	• English II, Algebra I, and Biology were administered operationally in both paper/pencil and online format (dual platform) starting in Fall 2008. These assessments consisted of both SR items and PE/WPs.
2009–2010	• English I, Algebra II, Geometry, Government, and American History were administered operationally in both paper/pencil and online format (dual platform) starting in Fall 2009. These assessments consisted of SR items only.
2010–2011	• PE/WPs were temporarily suspended from English II, Algebra I, and Biology starting in Summer 2010. • Assessments with SR items only (which include English I, Algebra II, Geometry, American History, and Government) were available in online format only.
2011–2012	• All assessments were administered online.
2012–2013	• PE/WPs were added back to English II, Algebra I, and Biology starting in Fall 2012.
2013–2014	• iPad and Chromebook administration were available for SR items in Summer 2013. • iPad and Chromebook administration were available for PE/WPs starting in Fall 2013.
2014–2015	• Physical Science was administered for the first time in Fall 2014. • Changes occurred for English I, English II, Algebra I, Algebra II, and Geometry, including revised blueprints, new test forms, and alignment of existing items to the Missouri Learning Standards. • Beginning in Fall 2014, English II, Algebra I, Algebra II, Biology, and Government are required and English I, Geometry, Physical Science, and American History are optional.
2015–2016	• A new Biology RSS table was used to score students for the Spring 2016 administration following a recalibration study.
2016–2017	• Student performance data revealed form comparability issues for the Algebra I and English II assessments. The results for these two tests were excluded from federal accountability.
2017–2018	• New Operational Forms based upon the new Missouri Learning Standards administered in Algebra I, Algebra II, Geometry, English I and English II. • A standard setting workshop was held to set new performance level standards for English I, English II, Algebra I, Algebra II, and Geometry after the first administration of new operational forms.
2018–2019	• New Operational Forms based upon the new Missouri Learning Standards administered in Biology and Physical Science. • A standard setting workshop was held to set new performance level standards for Biology and Physical Science after the first administration of new operational forms.

1.1.3. Current Administration of MO EOC Assessments

As the nine MO EOC assessments were administered in multiple forms in two current administrations (Fall 2018 and Spring 2019), it is helpful to clarify the coverage of this technical report. Table 1.2 presents the type of forms by content area and administration.

The two operational (OP) test forms (A and B) for English I, English II, Algebra I, Algebra II, and Geometry were newly developed for the 2017–2018 administration cycle. A standard setting workshop took place in July 2018 to set cut scores for performance levels. Another two operational test forms (C and D) for English and Mathematics contents were administered for the 2018–19 administration cycle. The two operational test forms (A and B) for Biology and one operational test form (A) for Physical Science were newly developed for the 2018–19 administration cycle. A standard setting workshop took place in 2019 to set cut scores for performance levels. The test forms for Government and American History were field-tested in stand-alone forms (SAFT) during the Spring 2018–19 operational administration.

The current technical report includes the assessments listed below:

- One or two operational test forms of English, Mathematics, and Science content areas in both Fall 2018 and Spring 2019
- Stand-alone field test forms of Government, American History, and Personal Finance in Spring 2019

Table 1.2. Type of Forms in Content Area by Administration

Content Area	Fall 2018	Spring 2019
English I	OP (A)	OP (C and D)
English II	OP (A)	OP (C and D)
Algebra I	OP (A)	OP (C and D)
Algebra II	OP (A)	OP (C and D)
Geometry	OP (A)	OP (C and D)
Biology	OP (A)	OP (A and B)
Physical Science	OP (A)	OP (A and B)
Government	--	SAFT
American History	--	SAFT

1.2. Purpose and Intended Use of MO EOC Test Scores

According to the *Standards for Educational and Psychological Testing* (AERA, APA, and NCME, 2014), Standard 1.1 states that:

> The test developer should set forth clearly how test scores are intended to be interpreted and consequently used. The population(s) for which a test is intended should be delimited clearly, and the construct or constructs that the test is intended to assess should be described clearly. (p. 23)

The Missouri State Board of Education identified the following purposes for the MO EOC assessments:

- Measures and reflects students’ mastery toward postsecondary readiness
- Identifies students’ strengths and weaknesses
- Communicates expectations for all students
- Serves as the basis for state and national accountability plans
- Evaluation of programs

The MO EOC assessments assess the Missouri Learning Standards and were created to meet the needs of Missouri districts, schools, teachers, and students while also meeting state and federal requirements. Evidence of students’ progress in meeting the Missouri Learning Standards is obtained from the MO EOC assessments. These assessments provide the data that DESE uses to inform students, parents, the public, and the state legislature about student performance to help make informed decisions about educational issues and drive student services throughout the state.

The intended interpretation of the MO EOC assessment scores is that the scores indicate students' progress toward mastering the Missouri Learning Standards. The interpretative argument involves the analysis of student performance in terms of individual achievement on the state standards and the conversion of these scores to performance levels (Kane, 2006). Student scores should facilitate proper interpretations while minimizing misinterpretations and unwarranted inferences. The MO EOC assessments incorporate the meaning of the test scores by anchoring the achievement level cut scores to known scale score values.

1.3. Validity Evidence and Validation Processes

Validity is the most fundamental consideration in educational and psychological testing. It refers to "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (AERA, APA, & NCME, 2014, p. 11). According to the *Standards for Educational and Psychological Testing*,

> Ultimately, the validity of an intended interpretation of test scores relies on all the available evidence relevant to the technical quality of a testing system…[this includes] evidence of careful test construction; adequate score reliability; appropriate test administration and scoring; accurate score scaling, equating, and standard setting; and careful attention to fairness for all test takers, as appropriate to the test interpretation in question. (p. 22)

The valid interpretation and appropriate use of MO EOC assessment scores are supported in a variety of ways. The validity evidence of score use and interpretation for any assessment stems from:

- the statement of the test's purpose and the intended use of the scores;
- the steps taken in designing the test; and
- the processes of developing the content of the test, consulting with stakeholders, communicating about the test to users, scoring and reporting, and conducting data analysis.

The documentation of each of these steps is a necessary piece of a comprehensive, defensible validity argument for the intended uses of the assessment scores. This document provides evidence necessary to assess the validity of the MO EOC assessment scores for their intended purposes.

The MO EOC assessments are part of an integrated program of testing, accountability, and curricular and instructional support. In reading this technical report, it is critical to remember that the assessment program does not exist in a vacuum; it is not just a test. It is one part of a complex network intended to help schools to improve student learning. The MO EOC assessments are an integrated program of testing and accountability, as well as curricular and instructional support. The assessments can only be evaluated properly within their full context.

This technical report provides details about the development and implementation of the MO EOC assessments. All information contained herein ultimately contributes to the argument for the validity of the interpretation and use of scores for their intended purposes. This section describes some of the aspects of validity evidence in this report.

1.3.1. Various Item Types

For 2018–2019, the English, Mathematics, and Science content area assessments contained selected-response (SR), technology enhanced (TE) items, and performance events/writing prompts (PE/WPs). A SR item presents students with a question followed by four or more response options. TE items include a variety of item types, such as drag and drop, free draw, text entry, extended text, line match, and graphing. PEs are open-ended items that require students to perform more complicated tasks. A PE measures depth of understanding and interpretative and analytical abilities in a format that allows for more than one approach to arrive at a correct response. The advantage of this item type is that it provides insight into a student's ability to apply knowledge and understanding in real-life situations. The WP, a special type of PE that appears in the English I and II assessments, is an open-ended item that requires students to demonstrate their writing proficiency.

1.3.2. Multiple Administrations

Testing for the MO EOC assessments is conducted during three state-designated windows each year for summer, fall, and spring. These tests are designed to be administered in approximately one testing period and are not strictly timed. The 2018–2019 MO EOC assessments were offered primarily in an online administration mode with Paper/Pencil, Braille, and Large Print forms available for students requiring accommodations.

1.3.3. Reporting the Results

The MO EOC assessment reports provide useful information for determining the performance of students in a particular school and classroom. These reports help identify students who are below Proficient in a particular content area so that the school may determine a course of action that will meet the students' specific needs. Districts may also use locally designed assessments aligned to the Missouri Learning Standards to provide more detailed information for each student in specific content areas.

Individual Student Reports (ISRs) and student raw scores are available to a district five business days after the close of their district testing window. Timely availability of score reports allows teachers the option to consider MO EOC assessment results in assigning course grades. ISRs are only available in an online format unless an order is placed by the district for paper reports.

1.4. Organizational Support

DESE coordinates the development and implementation of the MO EOC assessments. In addition to planning, scheduling, and directing all EOC activities, the staff is extensively involved in numerous test reviews, security, and quality assurance procedures. At the outset of the 2008 contract award, Riverside Publishing was the primary contractor working in partnership with Questar Assessment Inc. (Questar), the Assessment Resource Center (ARC), Internet Testing Systems (ITS), Bookette, and others. Beginning with the Summer 2011 administration, DESE contracted operational activities with Questar. Table 1.3 summarizes the main activities for each group involved with the 2018–2019 MO EOC administrations.

Table 1.3. Organizational Support

Group	Responsibilities
Questar Assessment Inc. (Questar)	• Provide program management, including primary contact with DESE; coordinate all meetings; handle all administrative costs/activities; generate all program management reports and status reports • Create and update the *Test Coordinator's Manual*, *Software Installation Guide*s, and other ancillary materials • Conduct psychometric analyses, reporting, linking/equating studies, and associated tasks • Provide all needed prepress work for program materials through camera-ready art • Produce all materials, including online, Paper/Pencil, Braille, and Large Print versions of the test, as well as online testing tools and content area-specific tutorials • Account for secure test books received after testing • Provide a direct customer service line, including technical support and general support to the program and customer interactions • Store materials after testing • Participate in and present at Technical Advisory Committee (TAC) meetings • Score all SR items and the PE/WPs • Produce and distribute all score reports and the *Guide for Interpreting Results* • Complete the technical report for DESE • Provide online enrollment and pre-ID system for use by Missouri districts • Provide online testing interface and online test administration site • Package and distribute materials • Barcode test books with security IDs
ACS	• Facilitated the standard setting workshop for the Biology and Physical Science EOC assessments in July 2019.
Districts	• Distribute materials to school buildings, track all secure materials, and promptly return all materials, including transcribed test forms, for scoring • Assist in the timely resolution of scoring alerts • Act as a liaison between Questar and buildings
School Buildings	• Administer tests, track all secure materials, and promptly return materials to districts for scoring
SeaChange Print Innovations	• Print Large Print versions
American Printing House for the Blind (APH)	• Print Braille versions

1.5. Chapter Summaries

Summaries of the information contained in the following chapters of this report are below.

Chapter 2: Test Content and Development

Chapter 2 provides the test blueprints with target point distributions and test specifications for the Fall 2018 and Spring 2019 administrations. Appendix A provides target and actual point distributions in blueprint. Information about item writing, content and bias reviews, test form construction, and statistical item review is also presented. The evidence is important to the content-related validity of the MO EOC assessment scores. This chapter also discusses principles of universal design and outlines the quality control processes employed throughout the test development process. Documentation of previous test designs can be found in the technical reports located on DESE's website at http://dese.mo.gov/college-career-readiness/assessment/assessment-technical-support-materials.

Chapter 3: Test Administration

Chapter 3 contains information about the administration of the MO EOC assessments. The chapter begins with testing windows and a description of students for whom the assessments are appropriate. Administration details are then summarized. This summary includes a description of how the materials are distributed and how Test Examiners are trained, as well as information about the organization of the assessments, preparation of students to take the assessments, and directions for administration. The chapter also includes information about the accommodations allowed on the MO EOC assessments and describes how materials are submitted for processing and scoring.

Chapter 4: Scoring

Chapter 4 covers the scoring processes for both the selected-response (SR) and performance events/writing prompts (PEs, WPs, and CRs) on the MO EOC assessments. It contains information on how Questar scored the MO EOC assessment, including the scoring training and qualification processes, scoring procedures, and monitoring for quality assurance. Finally, this chapter provides rater agreement for the Fall 2018 and Spring 2019 administrations. Information in this chapter provides evidence to support the validity and reliability of rater scores.

Chapter 5: Psychometric Analyses

Chapter 5 contains item-level analysis summary information and IRT based calibration, equating and scaling procedures. The classical item statistics include item difficulty and item discrimination indices for each content area for the Fall 2018 and Spring 2019 operational items. The results indicate that the MO EOC assessments have sound psychometric properties. The items measure achievement across a broad range of difficulty and most items are appropriately correlated with the total test score. The description of IRT based procedures of this chapter begins with an introduction to the item response theory (IRT) model used for building and maintaining of the scale of the MO EOC assessments. Next, the equating and scaling process for the Fall 2018 and Spring 2019 are provided. In particular, the equating procedures of English and Mathematics content area assessments described in detail. Finally, the raw-to-scale score (RSS) conversion tables for the operational forms are presented in Appendix E.

Chapter 6: Standard Setting and Cutpoint Validation

Chapter 6 summarizes the 2019 standard setting workshop that took place in July 2019. The chapter describes various features of the workshop including the external benchmark, description of the panel members, staffing, bookmark procedure, results, and the post-standard setting activities. The *Final Technical Report on the Standard-Setting Workshop for the Missouri Assessment Program* contains additional information on the 2019 standard setting workshop.

Chapter 7: Reliability and Construct-related Validity

Chapter 7 begins by defining reliability and providing an overview of reliability estimation techniques. Raw-score internal consistency reliability coefficients are presented for all students and for each demographic group. Classification accuracy and classification consistency statistics are also presented. The results indicate acceptable reliability and measurement precision. The validity evidence for the MO EOC assessments related to the internal structure of the assessments and other types of validity evidence proposed by the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) are followed in this chapter. It provides an argument supporting the validity of the MO EOC assessments for measuring Missouri students' mastery of the Missouri Learning Standards, for identifying students' strengths and weaknesses, for serving as a basis for evaluating accountability plans, and for program evaluation.

Chapter 8: Reporting and Results

Chapter 8 contains information about the reports Questar produced for the MO EOC assessments, including the Individual Student Report (ISR) and Student Score Label. A brief description of the state's data portal and reporting system is also included. The second part of this chapter provides descriptive statistics for raw scores and scale scores. Raw score statistics are summarized by test administration, content area, and cluster. Scale score statistics are summarized for each content area and are also broken down by gender and ethnicity as well as migrant, free and reduced lunch (FRL), limited English proficient (LEP), Title I, Individualized Education Program (IEP), and accommodation statuses.

Chapter 2: Test Content and Development

2.1. Introduction

On April 19, 2016, the Missouri State Board of Education approved new Missouri Learning Standards for ELA, Mathematics, Science, and Social Studies. The revised standards were implemented in the 2016–2017 school year. For English and Mathematics, these standards were assessed in 2017–2018. For Science, census field testing took place in 2017–2018 and operational testing began in 2018–2019. For Social Studies, census field testing took place in 2018–2019 and operational testing of the new standards will begin in 2019–2020.

New operational test forms were developed for English I, English II, Algebra I, Algebra II, Geometry, Biology, and Physical Science. Two core forms were administered for the English, Mathematics, and Science assessments. In Spring 2019, a stand-alone field test was administered for English writing prompts that will become operational on the remaining English core forms. The stand-alone field test forms were developed for American History and Government for the Spring 2019 administrations.

According to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014), "Important validity evidence can be obtained from an analysis of the relationship between the content of a test and the construct it is intended to measure" (p. 14). Accordingly, the descriptions of the test development procedures included in the MO EOC technical reports provide validity evidence of the MO EOC assessments. Documentation of test development from previous administrations, including the test designs, can be found in previous technical reports, located on DESE's website at http://dese.mo.gov/college-career-readiness/assessment/assessment-technical-support-materials.

2.2. MO EOC Content Standard

2.2.1. Content Validity

Baker and Linn (2002) suggests "Two questions are central in the evaluation of content aspects of validity. Is the definition of the content domain to be assessed adequate and appropriate? Does the test provide an adequate representation of the content domain the test is intended to measure?" (p. 6). The following sections help answer these two questions and address Standard 4.12[1], which specifically relates to the definition and development of test content.

2.2.2. Appropriateness of Content Definition

In 1993, the Missouri legislature passed the Outstanding Schools Act (Senate Bill 380) that required the State Board of Education to adopt challenging academic performance standards. These standards define the skills and competencies necessary for students to successfully advance through the public school system, prepare for postsecondary education and the workplace, and participate as citizens in a democratic society. The Missouri State Board of Education formally adopted the academic standards known as the Show-Me Standards in January 1996.

---

[1] **Standard 4.12:** Test developers should document the extent to which the content domain of a test represents the domain defined in the test specifications (AERA, APA, NCME, 2014, p. 89).

In addition to mandating the development of rigorous academic standards, the Outstanding Schools Act of 1993 required the development and implementation of a comprehensive, primarily performance-based assessment program to measure student proficiency in the knowledge, skills, and competencies identified in the Show-Me standards. Upon adoption of the standards in 1996, Missouri began developing the MAP.

In January 2007, the Missouri State Board of Education approved a plan to replace the MAP with end-of-course assessments for high school students. This transition occurred at the beginning of August 2008. The MO EOC assessments tested English II, Algebra I, and Biology. The remaining MO EOC assessments (English I, Algebra II, Geometry, Government, and American History) were added the following year. The intent was to provide MO EOC assessments that are an integral part of the statewide assessment system and, as such, are a logical extension of MAP Grade-Level assessments.

The Missouri State Board of Education approved new Missouri Learning Standards on April 19, 2016. These standards were implemented in 2016–2017. The MAP began assessing these standards in 2017–2018 for English and Mathematics. The new Science standards were assessed beginning in 2018–2019; the new Social Studies standards will be assessed beginning in 2019–2020.

2.2.3. Adequacy of Content Representation

The adequacy of the content representation of the MO EOC assessments is important because the tests must provide an indication of student progress toward achieving the knowledge and skills identified in the Missouri Learning Standards. The assessments must also fulfill the requirements of the Every Student Succeeds Act (ESSA).

The MO EOC assessments measure students' progress toward the Missouri Learning Standards, which are Missouri's content standards. Adequate representation of the content domains defined in the content standards is assured through the use of a test blueprint and a documented test construction process. The content standards were taken into consideration in the writing of all items. Evidence to support the content validity of the MO EOC assessments is provided in this Chapter through the documentation of the test specifications and blueprints, item-writing processes, and item-review processes. Specific efforts to ensure content validity are summarized below.

- Detailed test and item/passage development specifications were established; tests included sufficient numbers of items; and items were adequately distributed across content, levels of cognitive complexity, and difficulty.
- Qualified item writers were provided training.
- Missouri teachers were trained to create clear and simple instructions.
- Items were developed to include a wide array of contexts and cultures.
- Each newly developed item was first reviewed by content specialists and editors to ensure all items were aligned with the content standards. Appropriateness for the intended grade, depth of knowledge, graphics, grammar/punctuation, language demand, and distractor reasonableness were also considered.

- Missouri teachers from diverse ethnic and geographical backgrounds reviewed the items to ensure all items were accessible to as many students as possible.
- Content and bias review committees reviewed the items following specific criteria.

2.3. Test Blueprints

Test blueprints specify the relative percentage of items in each high-level content strand. Tables 2.1–2.10 provide the Fall 2018 and Spring 2019 test construction blueprints for English I, English II, Algebra I, Algebra II, Geometry, Government, and American History. Biology, Physical Science, and Personal Finance assessments were offered in Fall 2018 and Spring 2019. The test blueprints for the Fall 2018 and Spring 2019 administrations are presented for the operational tests only. Detailed test specifications that address coverage of each course-level expectation (CLE) are used to develop each test form.

Table 2.1. Test Construction Blueprint—English I

Content Strand	Point Range	Range of Emphasis
Reading Literary Text	15	30%
Reading Informational Texts	15	30%
Writing	20	40%
Total	50	100%

Table 2.2. Test Construction Blueprint—English II

Content Strand	Point Range	Range of Emphasis
Reading Literary Text	15	30%
Reading Informational Texts	15	30%
Writing	20	40%
Total	50	100%

Table 2.3. Test Construction Blueprint—Algebra I

Content Strand	Point Range	Range of Emphasis
Algebra	18-22	36-44%
Functions	18-22	36-44%
Number/Quantity and Statistics	8-12	16-24%
Total	50	100%

Table 2.4. Test Construction Blueprint—Algebra II

Content Strand	Point Range	Range of Emphasis
Algebra	25-28	50-56%
Functions	11-14	22-28%
Number/Quantity and Statistics	10-12	20-24%
Total	50	100%

Table 2.5. Test Construction Blueprint—Geometry

Content Strand	Point Range	Range of Emphasis
Congruence/Similarity, Coordinate Geometry & Circles	32-35	64-70%
Geometric Measurement & Modeling	6-10	12-20%
Stats and Probability	6-10	12-20%
Total	50	100%

Table 2.6. Test Construction Blueprint—Biology

Content Strand	Point Range	Range of Emphasis
From Molecules to Organisms: Structure and Process	11-15	22-30%
Ecosystems: Interactions, Energy, and Dynamics	8-12	16-24%
Heredity: Inheritance and Variation of Traits	11-15	22-30%
Biological Evolution: Unity and Diversity	11-15	22-30%
Earth and Human Activity	3-6	6-12%
Total	50	100%

Table 2.7. Test Construction Blueprint—Physical Science

Content Strand	Point Range	Range of Emphasis
Matter and Its Interactions	12-16	24-32%
Motion and Stability: Forces and Interactions	12-16	24-32%
Energy	12-16	24-32%
Earth and the Universe	6-9	12-18%
Total	50	100%

Table 2.8. Test Construction Blueprint—Government

Content Strand	Point Range	Range of Emphasis
Tools of Social Science Inquiry	10-12	20-24%
Historical Foundations	10-12	20-24%
Structure of Government	13-17	26-34%
Government in Action	13-17	26-34%
Total	50	100%

Table 2.9. Test Construction Blueprint—American History

Content Strand	Point Range	Range of Emphasis
Tools of Social Science Inquiry	8-10	16-20%
Re-Emerging America	8-10	16-20%
Emerging Globally	8-10	16-20%
The Great Depression and WWII	8-10	16-20%
The American Stage	8-10	16-20%
Contemporary America	8-10	16-20%
Total	50	100%

Table 2.10. Test Construction Blueprint—Personal Finance

Content Strand	Point Range	Range of Emphasis
Financial Decision Making/II. Earning Income	10-12	20-24%
III. Buying Goods and Services	10-12	20-24%
IV. Savings/V. Using Credit	15-18	30-36%
VI. Protecting and Insuring/ VII. Financial Investing	10-12	20-24%
Total	50	100%

2.4. Test Specifications

Standard 1.11[2] addresses the appropriateness of test content and its relationship to a solid validity argument. Additionally, Standard 4.2[3] defines test specifications and provides examples of the

[2] **Standard 1.11:** When the rationale for test score interpretation for a given use rests in part on the appropriateness of test content, the procedures followed in specifying and generating test content should be described and justified with reference to the intended population to be tested and the construct the test is intended to measure or the domain it is intended to represent. If the definition of the content sampled incorporates criteria such as importance, frequency, or criticality, these criteria should also be clearly explained and justified (AERA, APA, NCME, 2014, p. 26).

[3] **Standard 4.2:** In addition to describing intended uses of the test, the test specifications should define the content of the test, the proposed test length, the item formats, the desired psychometric properties of the test items and the test, and the ordering of items and sections. Test specifications should also specify the amount of time allowed for testing; directions for the test takers; procedures to be used for test administration, including permissible variations; any materials to be used; and scoring and reporting procedures. Specifications for computer-based tests should include a description of any hardware and software requirements (AERA, APA, NCME, 2014, p. 85–86).

type of information that should be included in a specifications document. The test specifications describe the content and format of the test and delineate the ideal number of items and points assessed for each standard.

While Tables 2.1–2.10 provide the target point distributions, Appendix A contains the actual point distributions. Details on the development and use of the test specification documents for previous MO EOC test forms can be found in previous technical reports on DESE's website at http://dese.mo.gov/college-career-readiness/assessment/assessment-technical-support-materials.

The following is an overview of the 2018–2019 test specifications:

- **English I**
  - The English I assessment measures student achievement in the following content strands:
    - Reading Literary Texts
    - Reading Informational Texts
    - Writing
  - The English I assessment has 40 OP points (consisting of SR and TE items), 12 FT items, and 1 WP with a score range of 0–2, 1–4, and 1–4 based on the three-part scoring guide, totaling 10 points.

- **English II**
  - The English II assessment measures student achievement in the following content strands:
    - Reading Literary Texts
    - Reading Informational Texts
    - Writing
  - The English II assessment has 40 OP points (consisting of SR and TE items), 12 FT items, and 1 WP with a score range of 0–2, 1–4, and 1–4 based on the three-part scoring guide, totaling 10 points.

- **Algebra I**
  - The Algebra I assessment measures student achievement in the following content strands:
    - Algebra
    - Functions
    - Number/Quantity and Statistics
  - Session 1 consists of SR and TE items, totaling 40 points, and Session 2 consists of one PE worth a total of 10 points. All items are aligned to the strands listed.
  - PEs are aligned to any of the strands listed, and while no set point value for each PE task is designated, the total must add up to 10 points.

- **Algebra II**
  - The Algebra II assessment measures student achievement in the following content strands:
    - Algebra

- Functions
- Number/Quantity and Statistics

  - Session 1 consists of SR and TE items, totaling 40 points, and Session 2 consists of one PE worth a total of 10 points. All items are aligned to the strands listed.
  - PEs are aligned to any of the strands listed, and while no set point value for each PE task is designated, the total must add up to 10 points.

- **Geometry**
  - The Geometry assessment measures student achievement in the following content strands:
    - Congruence/Similarity, Coordinate Geometry, & Circles
    - Geometric Measurement & Modeling
    - Statistics and Probability
  - Session 1 consists of SR and TE items, totaling 40 points, and Session 2 consists of one PE worth a total of 10 points. All items are aligned to the strands listed.
  - PEs are aligned to any of the strands listed, and while no set point value for each PE task is designated, the total must add up to 10 points.

- **Biology**
  - The Biology assessment measures student achievement in the following new content and process strands:
    - From Molecules to Organisms: Structure and Process
    - Ecosystems: Interactions, Energy, and Dynamics
    - Heredity: Inheritance and Variation of Traits
    - Biological Evolution: Unity and Diversity
    - Earth and Human Activity
  - The Biology assessment is comprised of 50 points (SR, TE, and Scenario Sets/PE).

- **Physical Science**
  - The Physical Science assessment measures student achievement in the following new content and process strands:
    - Matter and Its Interactions
    - Motion and Stability: Forces and Interactions
    - Energy
    - Earth and the Universe
  - The Physical Science assessment is comprised of 50 points (SR, TE, and Scenario Sets/PE).

- **Government**
  - The Government assessment measures a student’s ability to understand our history and participate in our civic life as citizens and consumers. The Government forms are worth 50 points and consist of SR, Constructed Response (CR), and TE items that are aligned to the following strands:
    - Tools of Social Science Inquiry
    - Historical Foundations

        - Structure of Government
        - Government in Action

- **American History**
    - The American History assessment measures a student's ability to understand U.S. history and participate in U.S. civic life as citizens and consumers. The American History forms are worth 50 points and consist of SR, CR, and TE items that are aligned to the following strands:
        - Tools of Social Science Inquiry
        - Re-Emerging America
        - Emerging Globally
        - The Great Depression and WWII
        - The American Stage
        - Contemporary America

- **Personal Finance**
    - The Personal Finance assessment measures a student's ability to understand and make decisions about real-world financial issues. The Personal Finance forms are worth 50 points and consist of SR items that are aligned to the following strands:
        - I. Financial Decision Making
        - II. Earning Income
        - III. Buying Goods and Services
        - IV. Savings
        - V. Using Credit
        - VI. Protecting and Insuring
        - VII. Financial Investing

2.5. Item Development

The process of constructing the tests that were administered in 2018–2019 is discussed in this section. Specifically, historical information regarding both item-development procedures and content coverage from Questar is presented. Content-related evidence of validity that supports test interpretation is presented in terms of how the MO EOC assessments were assembled.

Questar test development specialists created a detailed item and passage development plan based on the blueprints for each content area. The plans included the number of items necessary for each assessed course-level expectation (CLE) and an outline of the review process for developed items and passages. This process included internal Questar reviews, DESE item review, and a content and bias review by Missouri educators.

The forms for the English, Mathematics, Science, and Personal Finance assessments for Fall 2018 and Spring 2019 administrations were primarily constructed using items field tested in Spring 2018. During the process of building the forms for the operational test administrations, statistical characteristics (i.e., *p*-values and point-biserial correlations) were used to evaluate the items and test forms. The American History and Government stand-alone field test forms, which were administered in Spring 2019, included items from the 2018 Item Writer Workshop (IWW).

2.5.1. Item Writing

Missouri educators, DESE staff members, Regional Instructional Facilitators (curriculum and assessment specialists housed in each of Missouri's nine Regional Professional Development Centers), and Questar test development specialists created all the test items. For English II and English I PEs, permissioned passages were found by Questar passage searchers and approved by DESE staff and Missouri educators, and the corresponding writing prompts were written by item writers trained by Questar test development specialists and DESE staff. Requirements to be an item writer included experience in classroom teaching and expert content knowledge.

The Item Writing Workshop (IWW) for the Missouri EOC assessments in English, Mathematics, and Social Studies took place January 16–19, 2018, in Columbia, Missouri, and the IWW for Science took place January 22–25, 2018, in St. Louis, Missouri. The IWW for the Listening and Speaking items and Personal Finance assessments took place from June 11–15, 2018 in St. Louis, Missouri. The Listening and Speaking items are intended to be field-tested as part of the English I and English II assessments and will be used operationally in 2019-20. DESE invited participants from educational sites throughout Missouri for the purpose of authoring items aligned to the new Missouri Learning Standards for use as future field test items. The target number of items to be authored varied from course to course in keeping with the item development plans (IDP) and item writing (IW) assignments prepared by Questar. Program management, technical support, meeting logistics, oversight, as well as training and facilitation were led by Questar.

The workshop was held over two four-day periods and was conducted with 8 teacher participants per content area. Teacher participants were selected by DESE to represent school districts throughout Missouri. The content developed at the workshops was based on the updated Missouri Learning Standards and CLEs.

The English I and English II participants wrote Selected Response (SR), Technology Enhanced (TE), and Writing Prompt (WP) stand-alone items associated with the passages that had been approved prior to the item-writing workshops. At another IWW, educators wrote items to Listening passages for English I and II. The Algebra I, Algebra II, and Geometry participants wrote SR, TE, and PE items along with rubrics. Physical Science and Biology participants wrote SR and TE items along with scenario sets with rubrics. The American Government and History participants wrote SR and TE items, along with scenario sets, to the new standards that would be stand-alone field tested in 2018–2019.

During the IWWs, Questar test development specialists conducted training sessions with the item writers and provided instructions on avoiding bias and stereotyping of groups and individuals based on gender, race, ethnicity, religion, age, language, socioeconomic group, and disability. The IWW Training slide deck is available in Appendix B. Questar test development specialists also trained item writers to write items that adhere to the principles of universal design, making the items accessible to the widest range of students. For example, items and passages were written using clear and concise language, and all art, graphs, and tables were labeled and were not overly crowded with extraneous information. Instruction was also provided on developing items at particular cognitive levels based on Norman Webb's DOK levels.

Questar test development specialists trained item writers to enter content into Questar's electronic content management system. During training, each item writer wrote several items. Participants received feedback through the content management system where Questar test development specialists responded to teachers' items as they were submitted. As items were produced, they were continuously reviewed, revised, edited, and evaluated by Questar test development specialists and DESE staff. Item writers who generated high-quality work on or ahead of schedule were given additional assignments.

After a general session presentation and training, participants went into their content specific breakout groups. For most rooms, educators were grouped in pairs and IW assignments given to each pair of participants for completion. Rooms with odd numbers of participants were grouped in teams of two to three as needed. There were copies of the following materials in each room, in addition to any appropriate content specific materials (e.g., passages for authoring passage item sets or source materials):

- Item Writing Guide
- Missouri Learning Standards
- Content area item specifications
- Quick Notes for authoring items in the online authoring system
- Metadata notes for authoring item metadata in the online authoring system
- Guides for authoring each item type in the online authoring system
- IW Assignments

As items were written, they were tracked according to the item development plan. Questar kept records to maintain a workflow that generated items in assessment strands and CLEs as required by the test blueprint. All items and passages went through several rounds of internal reviews, including content and editorial reviews. Questar test development specialists reviewed each item with respect to alignment, clarity, and correspondence with item specifications.

2.5.2. Universal Design

Questar test development specialists were experienced in employing the principles of universal design in item development so that all students have equal access to the assessments. Questar included these principles when training Missouri teachers to write the items.
According to the NCEO Synthesis Report 44 (Thompson, Johnstone, & Thurlow, 2002), universally designed assessments have seven elements:

1. Inclusive assessment population
2. Precisely defined constructs
3. Accessible, nonbiased items
4. Amenable to accommodations
5. Simple, clear, and intuitive instructions and procedures
6. Maximum readability and comprehensibility
7. Maximum legibility

All items for the MO EOC assessments were developed with these elements in mind. Questar ensured the development of MO EOC items in accordance with these principles in the following manner:

- Items were developed to include a wide array of contexts and cultures. These item types may make students feel more included, increase motivation, and avoid bias.
- The test and item specifications served as a model for precisely defining the constructs that the tests would measure. These specifications indicated to the item writer, content reviewer, and test development specialists exactly what was to be measured. The item could assess a particular part of a standard or a combination of elements within a standard. The reviews served as a method for eliminating items that included assessment of knowledge outside the standard. For example, a Mathematics item should have nonmathematical vocabulary below grade level; otherwise the item might also be assessing reading ability, introducing construct-irrelevant variance.
- The review of items, which included Missouri teachers from diverse ethnic and geographic backgrounds, served to ensure that all items were accessible to as many students as possible.
- Questar staff members trained Missouri teachers to create clear and simple instructions so that students would have a clear understanding of the task needed to answer an item. Teacher review committees had an opportunity to review the instructions to ensure that they were appropriate for the grade levels and content areas. To ensure the appropriateness of the level of the vocabulary, Children's Writer's Word Book and EDL Core Vocabulary were employed by test developers and item review committees.
- Finally, items with text, art, tables, maps, and diagrams were constructed with maximum legibility.

Table 2.11 presents the number of item writers by content area. Table 2.12 presents the total number of items by item type that were generated during the Item Writer Workshop (IWW).

Table 2.11. Number of Item Writer Participants

Group	# Participants
English I	8
English II	8
Listening	12
Algebra I	8
Algebra II	8
Geometry	8
Biology	8
Physical Science (began with 8)	7
American History	8
Government	8
Personal Finance	5
Total	88

Table 2.12. Number of Items by Type at the End of the IWW

Content Area	Choice	Composite	Drag and Drop	Extended Text	Free Draw	Gap Match	Graphic Gap Match	Hot Text	Inline Choice	Line Graphing	Line Match	Match	Text Entry	Total
English I	142	3	9	--	--	6	--	24	6	--	1	13	--	204
English II	131	18	6	--	--	--	--	26	13	--	--	17	--	211
Algebra I	81	4	--	18	--	--	--	--	7	3	5	6	7	131
Algebra II	70	4	6	16	--	4	--	2	8	1	11	4	2	128
Geometry	62	16	1	11	1	--	5	--	20	1	18	3	29	167
Biology	139	10	3	4	--	--	2	--	14	--	--	10	1	183
Physical Science	117	38	12	38	--	--	7	5	38	--	1	16	13	284
Government	81	26	1	13	--	--	7	6	2	--	9	13	--	158
American History	180	2	19	24	--	--	--	13	--	--	6	13	--	260
Listening	239	--	--	--	--	--	--	--	--	--	--	--	--	239
Personal Finance	142	--	3	--	--	--	--	--	--	--	--	1	--	146

2.5.3. Content and Bias Review Process

Standard 4.8[4] addresses the importance of item review by an examination of the item statistics and the use of expert panels of judges. This section details the steps taken to ensure that the items chosen for the operational forms of the MO EOC assessments were of high technical quality and were free from bias. The Content and Bias Review in English I, English II, Algebra I, Algebra II, Geometry, American History, Government, Biology, and Physical Science took place June 11, 2018 through June 15, 2018 in St. Louis, Missouri; the Content and Bias Review for writing prompts, the new Listening items, and Personal Finance assessments took place October 2, 2018 through October 4, 2018 in Columbus, Missouri. The content review committees included DESE staff, Missouri educators from around the state, Regional Instructional Facilitators, and Questar staff.

The content and bias review committees reviewed SR items and PE/WPs using the following criteria:

- Overall quality and syntactical clarity
- Content coverage and content appropriateness
- Alignment to the specified CLE
- Appropriate contexts

[4] **Standard 4.8:** The test review process should include empirical analyses and/or the use of expert judges to review items and scoring criteria. When expert judges are used, their qualifications, relevant experiences, and demographic characteristics should be documented, along with instructions and training in the item review process that the judges receive (AERA, APA, NCME, 2014, p. 88).

- One clearly correct answer and plausible distractors for SR items
- Free from bias or any racial, socioeconomic, gender, or other sensitivity issues

The bias review committee was held separately from the content review committee and focused on reviewing items on the last criterion above. Suggestions from the bias review committee were then shared with the content review committee for their review and a determination on how to incorporate the edits.

Bias reviewers assessed each item for sensitivity of item context based on the guidelines delineated below, including, but not limited to, familiarity of language, possible stereotypes in context, and any potential advantages or disadvantages the context or content of an item might provide to a student or group of students. Guidelines for sensitivity reviews were as follows:

- Ensure that language has the same basic semantic content for all students regardless of race, gender, ethnicity, age, sexual orientation, or physical or mental condition.
- Consider the denotative and connotative meaning of words, expressions, images, and symbols.
- Avoid items or materials that might evoke negative or potentially inflammatory associations on the part of students.
- Consider the fairness of items and supporting materials:
    - Include females and males, and reflect nontraditional and traditional roles, relationships, and traits and occupations.
    - Present women, very young and elderly individuals, individuals with disabilities, a range of religious, ethnic, and racial minorities in roles of diverse status and power, conventional and unconventional.
- Ensure that there is no stereotyping.
- Represent the multiculturalism and diversity of our schools, nation, and world.
- Consider the possibility of sensitivity toward particular topics, which may interfere with students' ability to address item directives. Topics often regarded as sensitive include the following:
    - Death/suicide
    - Extreme illness
    - Violence/terrorism
    - Religion
    - Sex/birth control/pregnancy
    - Drugs/alcohol
    - Bigotry/bias
    - Homelessness
    - Family dysfunction
- Avoid creating situations in which students are asked to, or feel compelled to, divulge personal information (e.g., religious, social, and economic disclosures).

Before reviewing the items, a group training session was held with all committee members. Questar presented a PowerPoint that described the MO EOC program, the test development process, and the content and bias review procedures. After the large-group session, the committee members went to their respective break-out rooms to discuss the week's activities in more detail. The committee members were provided with copies of the Missouri Learning Standards and item specifications for the content area they were to review. Each Questar content facilitator reviewed these documents with the committee and answered any questions. The committee members were given the following checklists that could be referenced throughout the review process:

For all items:

- ☐ Does the item assess the assigned Missouri Learning Standards?
- ☐ Is the item clear, concise, and complete?
- ☐ Does the item contain accurate and sufficient content information?
- ☐ Is the item grade-level appropriate, and are the vocabulary and syntax appropriate for the students at the intended grade? (Reference the EDL Core Vocabularies.)
- ☐ Is the item fair to all students and free of bias and sensitivity issues?
- ☐ Does the item have correct punctuation, and is it grammatically correct?
- ☐ Is the item free from spelling and typographical errors?
- ☐ Is clueing avoided within an item stem and options, as well as among items?
- ☐ Does the item stand alone? (The answer to one item should not be dependent on the content of another item.)
- ☐ Are the equations, tables, charts, graphs, and other art clear, accurate, and necessary?
- ☐ Does the item have only one correct answer? (except in multi-select items).
- ☐ Does the item have unique, plausible distractors containing common errors students would make?
- ☐ Options are parallel and balanced, and outliers (e.g., use of key words from the stem, negatives, proper nouns, numerals) are avoided.
- ☐ Do all distractors contain clear rationale statements?
- ☐ Is the item free from absolutes (“none of the above,” “all of the above”) as options and free from the use of negatives (“not,” “none,” “except”) in the stem?
- ☐ Does the item avoid repeating words from the stem in the options?
- ☐ Does the item pose a single problem (although the solution may require more than one step)?
- ☐ Options are plausible and passage-based (for ELA).
- ☐ Options are grammatically and syntactically compatible with the stem.
- ☐ Options are stacked short-to-long or long-to-short.
- ☐ Direct quote options from the passage are ordered as they appear in the passage.

Technology checklist:

- ☐ The use of technology is justified (i.e., the item allows the student to respond in a way that is not possible or is not efficient via a traditional multiple-choice item).
- ☐ The technological aspects of the item do not introduce unnecessary demands on students.
- ☐ The standard that the item assesses lends itself well to the use of the format.

For PE/WPs:

- ☐ Does the item assess the assigned Missouri Learning Standards?
- ☐ Does the item clearly specify how the student should respond?

- ☐ Does the item allow for a variety of acceptable responses for the student to get full credit?
- ☐ Is the item grade-level appropriate, and are the vocabulary and syntax appropriate for the students at the intended grade? (Reference the EDL Core Vocabularies.)
- ☐ Is the item rich enough to elicit an appropriate range of responses covering all possible score points?
- ☐ Is the item fair to all students and free of bias and sensitivity issues?
- ☐ Does the rubric clearly define an acceptable answer or answers at each score point level?

Missouri educators participated in the review process for each content area. The number of participants by content area is presented in Table 2.13. The committee members read and reviewed each item. Discussions were held about whether the items met the criteria listed above. The committees then rejected or revised any items they deemed unsatisfactory. If there was disagreement about how to proceed with an item, the Questar facilitator polled the group and followed the direction of the majority. Table 2.14 shows the number of items reviewed and accepted in 2018. The accepted items were placed in a pool of items from which the 2018–2019 Government and American History stand-alone field test forms were built. The accepted items for the English, Mathematics, Science, and Personal Finance were placed on operational forms in the 2018–2019 administrations.

All item review sessions were held in secure meeting rooms, and all materials were confidential. Committee members were required to sign confidentiality agreements so that the integrity of the test content was not compromised. Although educators were encouraged to share information with their colleagues about the process of the item review, they were made fully aware of the expectation that any information about specific items and passages was to remain secure and confidential.

Table 2.13 Number of Content Bias Review Participants by Content Area

Group	# Participants
English I	8
English I and II SpEd representative	1
English I Speaking and Listening	10
English II	8
English II Speaking and Listening	9
Algebra I	8
Algebra II	8
Geometry	8
Biology	8
Physical Science	7
Government	8
American History	8
Personal Finance	7
Total	**98**

Note. SpEd stands for Special Education.

Table 2.14. 2018 Content/Bias Item Review Acceptance Rates

Content Area	Total #Items Presented for Review	#Items Accepted (as-is or with edits)	Acceptance Rate (items accepted as-is or with edits)
English I	146	144	98.6%
English I-WP	10	10	100%
English I-Listening	123	108	87.8%
English II	140	138	98.6%
English II-WP	10	9	90%
English II-Listening	118	107	90.7%
Algebra I	117	117	100%
Algebra II	114	112	98.2%
Geometry	124	122	98.4%
Biology	177	154	87%
Physical Science	171	151	88.30%
Government	136	123	90.4%
Am. History	213	190	89.2%
Personal Finance	271	271	100%

2.6. Pilot and Field Testing

2.6.1. Field-test Selection and Administration

The items accepted at the content/bias review were used to build the embedded and stand-alone field test forms administered in Fall 2018 and Spring 2019. Each fall ELA and Mathematics forms used 12 embedded FT items while each spring form included 10 embedded field test forms. There were 32 core forms (including accommodated forms) for the spring and fall administrations. Field-test items were selected so that each form met the established operational blueprint requirements for content coverage as closely as possible as shown in Table 2.15.

Table 2.15. Field Test Forms

Content Area	Form 1 (Core 1/FT) TTS					Form 2 (Core 2/FT)				
	OP	WP	FT	PE	FT PE	OP	WP	FT	PE	FT PE
English I	40	1	12	--	--	40	1	12	--	--
English II	40	1	12	--	--	40	1	12	--	--
Algebra I	40**	--	9	1	1-2	40**	--	9	1	1-2
Algebra II	40**	--	9	1	1-2	40**	--	9	1	1-2
Geometry	40**	--	9	1	1-2	40**	--	9	1	1-2
Biology	--	--	40-50	--	--	--	--	40-60	--	--
Physical Science	--	--	40-50	--	--	--	--	40-60	--	--
Government	40	--	10	--	--	--	--	--	--	--
American History	40	--	10*	--	--	--	--	--	--	--
Personal Finance	40	--	10	--	--	--	--	--	--	--

Note. * indicates placeholder items; ** 40 OP points, not items
American History/Government were stand-alone field test forms.

2.6.2. Classical Item Analyses

The statistics computed for the field test items are described below. The *p*-value and the item-test correlation indicate the item's difficulty and discrimination, respectively, and differential item functioning (DIF) was used to identify items that are potentially unfair.

- Item difficulty (*p*-value)
  - The *p*-value indicates how easy or hard an item is and is bound by 0 and 1. For the items worth one point, the *p*-value is the proportion of students who answered an item correctly. For items worth more than one point, the *p*-value is the average item score divided by the total possible points. The following was also presented:
    - The percentage of students choosing each option for the multiple-choice (MC) items;
    - The percentage of students obtaining each score point for other item types.

- Item discrimination (item-test correlation)
  - The correlation indicates how well an item distinguishes between low- and high-performing students and ranges from –1 to +1.
  - The correlation for each item was computed using students' scores on the field test item and students' total operational test score. Since all the items of interest are field test items, the operational test score did not include the item of interest.
  - An item with a high correlation indicates that students who do well on the total test tend to answer the item correctly and students who do poorly on the total test tend to answer the item incorrectly.
  - The point-biserial correlation, a special case of the Pearson product-moment correlation, was used for any item worth one point, like the MC items. The Pearson product-moment correlation was used for the items worth more than one point.
  - Correlations were computed for the distractors of the MC items and for each score point for the other item types.

- Differential item functioning (DIF)
  - DIF compares item performance between two groups of students who are matched on overall ability. It is expected that students who have comparable knowledge as measured by the test should perform similarly on the item.
    - DIF occurs when students from two different subgroups perform substantially different on an item but perform similarly on the test as a whole.
    - The presence of DIF does not necessarily indicate bias. Sometimes the knowledge or skill assessed by an item happens to be more common in one group than in another group. The presence of DIF should be considered as evidence that bears further investigation.

- Items were classified into three categories. Items classified as category C DIF were flagged.
  - A = negligible DIF
  - B = slight to moderate DIF
  - C = moderate to large DIF
- DIF was performed using the Mantel-Haenszel (MH) (1959) procedure for dichotomous items and the standardized mean difference (SMD) (Dorans, Schmitt, and Bleistein, 1992) for polytomous items.
- For the MH procedure, the odds ratio was converted to the delta metric, and the Educational Testing Service categorization was applied to flag the degree of DIF effects (Dorans & Holland, 1992). If the absolute value of delta was smaller than 1.00, the item was categorized as A. If the absolute value of delta was larger than or equal to 1.50, the item was classified as C. Otherwise, items were categorized as B.
- For the Winsteps DIF analyses, the level of DIF was determined by the absolute logit value of the DIF contrast. Absolute logit values less than 0.43 were classified as A, greater than or equal to 0.64 were classified as C, and between 0.43 and 0.63, inclusively, were classified as B (Linacre, 2015).
- Group comparisons were Male vs. Female, White vs. Hispanic, and White vs. African American.

DIF analyses were performed when there was a minimum of 200 students in the focal group.

2.6.3. Statistical Data Review

After completion of the 2018–2019 assessment windows, Questar test development specialists and psychometricians reviewed the statistical characteristics of the items. Questar used classical item statistics, including *n*-counts, *p*-values, percentage choosing each response option, point-biserial correlations, and differential item functioning (DIF) analysis.

During the data review, Questar Research and Assessment Development staff and DESE staff reviewed student performance on the Spring 2019 field test items for all EOC courses except for Government and American History. Items were reviewed regarding their statistical characteristics. Item reviewers from DESE and Questar were provided with the following information:

- Form
- Position
- Item as it appeared in the printed books
- Item alignment to the Missouri Learning Standards
- The *p*-value of the correct answer and percentage of students who selected each distractor (for SR items only)
- Mean and SD of item score (for PE/WPs only)
- Point-biserial correlation of correct response and point-biserial for each distractor (for SR items only)

- Total number of students who attempted to answer each question
- DIF using the Mantel-Haenszel (MH) (1959) procedure and the Educational Testing Service (ETS) classification (for SR items only)

Questar and DESE staff reviewed items that were flagged because of statistics that fell outside the parameters determined by the Questar Research staff. Table 2.16 contains the guidelines that were used for data review.

Table 2.16. Criteria for Flagged Items

Item Flagging Criteria	Indicates
If *p*-value of keyed response < 0.35	Difficult item
If *p*-value of keyed response > 0.95	Easy item
If *p*-value of keyed response < *p*-value of distractor	Possible miskey
If *p*-value of distractor > 0.35	Possible second correct option
If point-biserial of keyed response < 0.20	Poorly discriminating item
If point-biserial of a distractor is > 0.00	Possible second correct option
If ETS classification is B or C (from DIF analysis)	Possible bias in item

Each flagged item was reviewed; Questar and DESE then decided whether the item should be accepted or rejected. The review included items flagged with moderate to severe DIF (an ETS classification of B or C). Table 2.17 provides the number of items field tested and the non-flagged and flagged items by content area.

Table 2.17. Number of Flagged Items by Content Area

Content Area	No Flag	Flag	Total
English I	110	58	168
English II	119	58	177
Algebra I	63	43	106
Algebra II	59	46	105
Geometry	37	67	104
Biology	100	37	137
Physical Science	19	41	60
Total	507	350	857

Table 2.18 presents the number of items flagged for each criterion. Items were most frequently flagged for poor discrimination, high item difficulty, and a positive correlation for a distractor. No item was flagged for being too easy (*p*-value > 0.95). Across the content areas, 43 percent of items had no flags (*n*=180), 24 percent had one flag (*n*=99), 19 percent had two flags (*n*=82), 6 percent had three flags (*n*=24), 8 percent had four flags (*n*=34), and less than one percent had 5 flags (*n*=2).

Table 2.18. Items Flagged by Criterion

Content Area	FT Items	Low *p*-Value	High *p*-Value	Popular Distractor	Low Correlation	Distractor Correlation	C DIF
English I	168	12	3	23	33	34	12
English II	177	14	1	17	36	36	1
Algebra I	106	27	1	13	12	10	1
Algebra II	105	23	3	9	7	11	7
Geometry	104	43	0	21	20	19	9
Biology	137	22	2	13	12	13	4
Physical Science	60	18	0	8	29	18	5
Total	857	159	10	104	149	141	39

A flagged item was accepted if the review team determined that the item was strong and tested students on content they were expected to know. Accepted items were then made available in the pool of items that could be used to create the operational forms. Items the review team felt were biased or inappropriate for the MO EOC assessments were rejected. Rejected items were removed from the item pool, making them invalid for the MO EOC assessments.

2.6.4. Results

Table 2.19 provides the data review meeting results. The numbers of items that were field tested, flagging status, and, if flagged, rating is presented by content area and reporting category. Out of 857 items that were field tested, 789 were accepted (92.1%), 23 were revised (2.7%), and 45 were rejected (5.3%). Revised items will be re-field tested.

Table 2.19. Data Review Rating by Reporting Category

Content Area	Reporting Category	FT Items	Flag Status		Rating		
			No Flag	Flag	Accept	Reject	Revise
English I	Reading Informational Texts	38	23	15	34	4	0
	Reading Literary Texts	39	20	19	33	6	0
	Speaking/Listening	75	57	18	68	7	0
	Writing	16	10	6	15	0	1
	Total	**168**	**110**	**58**	**150**	**17**	**1**
English II	Reading Informational Texts	38	22	16	37	1	0
	Reading Literary Texts	37	23	14	34	3	0
	Speaking/Listening	86	67	19	81	5	0
	Writing	16	7	9	15	0	1
	Total	**177**	**119**	**58**	**167**	**9**	**1**
Algebra I	Algebra	42	28	14	41	0	1
	Functions	44	26	18	42	2	0
	Number/Quantity and Statistics	20	9	11	20	0	0
	Total	**106**	**63**	**43**	**103**	**2**	**1**
Algebra II	Algebra	49	35	14	49	0	0
	Functions	38	17	21	37	1	0
	Number/Quantity and Statistics	18	7	11	17	1	0
	Total	**105**	**59**	**46**	**103**	**2**	**0**

Content Area	Reporting Category	FT Items	Flag Status		Rating		
			No Flag	Flag	Accept	Reject	Revise
Geometry	Congruence/Similarity, Coordinate Geometry, and Circles	61	25	36	55	3	3
	Statistics and Probability	20	3	17	15	4	1
	Geometric Measurement and Modeling	23	9	14	20	1	2
	Total	**104**	**37**	**67**	**90**	**8**	**6**
Biology	From Molecules to Organisms: Structure and Process	30	20	10	29	0	1
	Ecosystems: Interactions, Energy, and Dynamics	30	24	6	26	2	2
	Heredity: Inheritance and Variation of Traits	34	27	7	32	0	2
	Biological Evolution: Unity and Diversity	33	21	12	29	1	3
	Earth and Human Activity	10	8	2	10	0	0
	Total	**137**	**100**	**37**	**126**	**3**	**8**
Physical Science	Matter and Its Interactions	18	4	14	14	2	2
	Motion and Stability: Forces and Interactions	17	4	13	14	1	2
	Energy	18	9	9	15	1	2
	Earth and the Universe	7	2	5	7	0	0
	Total	**60**	**19**	**41**	**50**	**4**	**6**
	Grand Total	**857**	**507**	**350**	**789**	**45**	**23**

2.7. Form Construction

2.7.1. Online Form Construction

In 2010–2011, Missouri began moving toward a full implementation of online administration of all MO EOC assessments. To assist in a smooth transition to online administration of all MO EOC assessments without interruption of data trends, Questar completed an online comparability study (see the *2013–2014 MO EOC Technical Report* for the full report). Based on the results of the study, the MO TAC reached a consensus that the move from Paper/Pencil to online administration would not affect student performance. As such, all 2018–19 EOC assessments (with the exception of the Paper/Pencil, Braille, and Large Print test forms for students needing such accommodations) are available on Questar's online delivery platform, Nextera®. More information on the current online test administration can be found in Chapter 3.

2.7.2. Quality Control for Form Construction

Checklists and quality control procedures accompanied each stage of form construction. A list of some quality control procedures used during the assembly of the MO EOC assessment forms is below:

- Construct forms based on all content requirements noted in the test blueprint and test specifications.
- Verify correct number of items per standard or reporting category based on test blueprint.
- Review items to ensure a wide sampling of the knowledge and skills being measured.
- Ensure that all items have been through the appropriate review procedures and are approved for use by DESE.

- Check for a variety of item topics, equal distribution of males and females, ethnicities, etc.
- Verify appropriate portions of items with and without artwork.
- Check for clueing across all items on each form.
- Verify equal or nearly equal distribution of answer choices for SR items.
- Ensure that the test meets the required statistical specifications (i.e., that as many items as possible have *p*-values between 0.35 and 0.90 and as many items as possible have point-biserial correlations above 0.20).
- Consider any statistical flags or problems.
- Check statistics to ensure that the collection of items on a given form yields an overall difficulty that falls within the specified range.
- Verify that items have not been released to the public.
- Verify correct answer key for each item.
- Perform content review of form (senior staff).
- Perform statistical review of form (psychometrician/statistician).
- Send form to DESE for review and approval.

2.7.3. Braille and Large Print Versions

Beyond employing the principles of universal design, all operational assessments were offered in Paper/Pencil (for students requiring a paper form of the assessment), Braille, and Large Print versions for visually impaired students taking the MO EOC assessments. To accommodate these students, a Braille and a Large Print paper version of the test were available. Once the Braille and Large Print forms were created for each assessment, reviews were held with DESE educators who had specialized training in working with visually impaired students.

The teachers consulted the Large Print and Braille Style Guide, which was also used during form composition, and relied on their own expertise to determine whether changes to directions, passages, or items were needed, or whether items should be omitted. Questar's Braille vendor (APH) also reviewed the forms and made recommendations based on how items, passages, and directions would be transcribed to Braille.

Questar and DESE reviewed the recommendations from all of these sources to determine if any required items to be omitted to accommodate the three versions. Table 2.20 below shows the breakdown. Items omitted from the operational assessment were items that would not Braille appropriately. The items may be TE items or items with art. Students taking the Braille form were given credit for these items. The embedded field test (EFT) items were eliminated from both the Braille and Large Print versions of these forms due to the irregular testing conditions and the small sample sizes for these groups. For 2018–19, a single Braille and Large Print test version was used for all MO EOC assessments.

Table 2.20. Accommodated Forms

Content Area	Accom Form 1 PP					Accom Form 1 LP					Accom Form 1 BR				
	OP	WP	PE	Omits	FT Slots	OP	WP	PE	Omits	FT Slots	OP	WP	PE	Omits	FT Slots
English I	40	1	--	12	--	40	1	--	12	--	40	1	--	12	--
English II	40	1	--	12	--	40	1	--	12	--	39	1	--	12	--
Algebra I	40	--	1	10	--	40	--	1	10	--	40	--	1	10	--
Algebra II	40	--	1	10	--	40	--	1	10	--	40	--	1	10	--
Geometry	40	--	1	10	--	40	--	1	10	--	40	--	1	10	--
Biology	41	--	--	--	--	41	--	--	--	--	--	41	--	--	--
Physical Science	46	--	--	--	--	46	--	--	--	--	--	46	--	--	--
Government	40	--	--	--	10	40	--	--	10	--	40	--	--	10	--
American History	40	--	--	--	10	40	--	--	10	--	40	--	--	10	--
Personal Finance	40	--	--	--	10	--	--	--	--	--	--	--	--	--	--

Note. Biology and Physical Science were SAFT forms, so items were chosen from the SAFT forms that would easily convert to PP, LP, and Braille, hence no omits.

2.8. Summary

The MO EOC assessments provide an indication of student progress toward achieving the knowledge and skills identified in the Missouri Learning Standards. Just as the content standards guided the item development and selection process, the consideration of content played an equally important role in form development. Form development required a balance of both content coverage and item difficulty. As items were selected for inclusion on particular forms, every effort was made to balance the content coverage to ensure the items aligned to the content standards being assessed while simultaneously considering the overall difficulty of the forms.

Chapter 3: Test Administration

3.1. Introduction

This chapter contains information about DESE and Questar's processes that ensure the standardized administration of the MO EOC assessments. The *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) state, "For tests designed to assess the test taker's knowledge, skills, abilities, or other personal characteristics, standardization helps to ensure that all test takers have the same opportunity to demonstrate their competencies" (p. 111). In other words, attention to the details of information dissemination, Test Examiner training, accommodations and modifications, and test security help ensure that students taking the MO EOC assessments in different locations and under different circumstances have comparable opportunities for success.

The EOC *Test Coordinator's Manual* contains detailed information about the testing guidelines, materials handling, and standardized administration instructions for the MO EOC assessments. While this manual is not included here, much of the information contained in this chapter can be found in it.

Questar uses its online assessment platform to manage and deliver the MO EOC Online assessments. This platform has two components:

- Student Test Delivery – The online testing student client is a small-footprint, secure browser application that is downloaded to the students' workstations to allow uninterrupted testing and failsafe protection of student responses in the event of a connection loss.
- Administration and Reporting System – The online testing system administration system is a web application that allows districts, schools, and teachers/proctors to manage their students and assessments.

For the MO EOC assessments, 2011–2012 was the first year in which districts were required to use an online delivery format unless a Paper/Pencil, Braille, or Large Print edition was required for a student as indicated in the student's Individualized Education Program (IEP) and marked as an accommodation on the online test administration site. The *Test Coordinator's Manual* contains information specific to the registration for and administration of the MO EOC assessments. This process was continued for 2018–2019.

3.2. Testing Calendar

Table 3.1 displays the 2018–2019 MO EOC testing windows. Each MO EOC assessment is tailored to each EOC content area and is designed to be administered when a student has completed the content defined for that course. Multiple testing windows allow school districts the flexibility to schedule MO EOC testing as close as possible to the end of each course so that they can provide students the greatest opportunity to demonstrate proficiency in the course content.

Table 3.1. Testing Windows

Test Period	Dates
Fall 2018	October 1, 2018 – January 18, 2019
Spring 2019	February 18, 2019 – May 31, 2019

Districts can offer EOC course content in any grade and in a variety of configurations. Although many districts offer EOC course content within a course bearing the same name, EOC course content can also be embedded within a course or across several courses. MO EOC assessments are administered according to a "right test, right time" philosophy when students have completed the appropriate content.

3.3. Students for Whom the MO EOC Assessments are Appropriate

The responsibility and authority for testing students in the MO EOC assessments at the appropriate time in the course of instruction belongs to the local district. The MO EOC assessments are based on Missouri Learning Standards rather than on GLEs. Therefore, when the content of the Missouri Learning Standards is covered in the local school district's curriculum, the test may be administered regardless of student grade level or course name.

3.3.1. Students with Individualized Education Programs (IEPs)

A student with disabilities, as classified under the Individuals with Disabilities Education Act (IDEA), has an IEP that, in part, governs whether a particular assessment is appropriate for the student. In the case of the MO EOC assessments, decisions about whether a student with a disability will participate in the assessments are made by the student's IEP team and are documented in the IEP. All students must take required MO EOC assessments. If, however, a student's disability qualifies him or her to take the MAP-Alternate assessment (MAP-A) for students with severe cognitive disabilities, that student will not participate in the MO EOC assessments.

3.3.2. Students with Individual Accommodation Programs

Students with Individual Accommodation Programs (IAPs) are considered disabled under Section 504 of the 1973 Rehabilitation Act. These students are not served under IDEA and are not documented with a particular designation for the MO EOC assessment. However, professionals who are knowledgeable about a student's disability and educational needs should make accommodation decisions for the student as they would for a student with an IEP.

3.3.3. English Language Learner (ELL) Students

Students who have been in the United States for 12 cumulative months or less since school age at the time of test administration may be exempted by the local school district from taking the English I and English II assessments. The students must, however, participate in other required MO EOC assessments, although their scores do not count for school accountability purposes. The other MO EOC assessments that all students, including ELL students, are required to take are Algebra I, Biology, and Government.

3.4. Students for Whom a School or District is Accountable

For accountability purposes, Missouri must include the results for any student who is eligible to take the MO EOC assessments and has been enrolled for at least one full academic year in a school (for school accountability) or district (for district accountability) without transferring out of the building or district for a significant period of time and re-enrolling. A full academic year is defined as the last Wednesday in September through the MO EOC assessment administration. A significant period of time is considered "one more than half of the eligible days between the last Wednesday in September and the test administration." DESE obtains enrollment information from the Missouri Student Information System (MOSIS) data that are reported by school districts. This rule applies to the building and district summary levels independently. For example, a student who is coded as "In building less than a year" but was in the district a full academic year is excluded from the building totals but is included in the district totals.

3.5. Dissemination of Testing Materials and Information

All test administration information, including the *Test Coordinator's Manual* and training webinars, were posted to the online test administration site for District Test Coordinators, School Test Coordinators, Examiners, and Information Technology Coordinators. One week prior to the start of the testing window, Questar distributed all password information for the online system by e-mail to district and school level users participating in the current EOC administration. Districts had the opportunity to order the Braille and Large Print editions of the assessment from Questar. The District Test Coordinator downloaded and printed the accommodated Paper/Pencil edition through the online administration site, as needed for students in the district. The District Test Coordinator was responsible for inventorying all Paper/Pencil materials, as well as disseminating the online test information to the test administrators. The District Test Coordinator was also responsible for answering all district questions about test procedures and the online assessment platform. If the District Test Coordinator needed assistance with a question, he/she could contact Questar's Missouri Customer Service through the designated phone number and/or e-mail address.

3.6. District and Test Examiner Training

Both Questar and DESE were responsible for training the district staff on EOC test administration. Questar and DESE provided training webinars, scripts, and PowerPoint presentations on the *Test Coordinator's Manual*, state procedures, and general testing issues. These training resources were available both on the DESE website and on the online test administration site. Appendix H contains the 2018–2019 training PowerPoint presentations for the MO EOC assessments.

Questar provided both onsite and recorded trainings on the online assessment platform. Questar training contained proprietary information and was only available on the test administration site. All Test Coordinators and Test Examiners were to view these standardized trainings prior to test administration. The District Test Coordinator was allowed to provide supplemental training on local issues (e.g., schedules). Both DESE and Questar were available to answer any questions the districts may have had about the MO EOC assessment administration.

3.7. Test Security

3.7.1. Summary

The MO EOC assessment test books (Paper/Pencil, Large Print, and Braille) and online assessments were secure. Test Coordinators were instructed to keep the materials in a locked room or cabinet at all times when not in use. No testing materials could be photocopied, duplicated, scanned, or made accessible to personnel who were not responsible for testing. Additionally, written or oral discussion of specific MO EOC assessment items breaches the security and integrity of the test. In accordance with the Standards, the *Test Coordinator's Manual* contained explicit instructions about test security for Test Coordinators and Test Examiners.[5]

Standardized training was required for all District and School Test Coordinators, Examiners, translators, proctors, and any district staff who had responsibilities in testing. Each test book shipped to the district or downloaded and printed by the district contained secure barcode information for tracking purposes. Questar used this information to ensure that districts used the materials assigned to them for testing and returned all of their secure materials after the completion of testing. The Paper/Pencil forms included a barcode on each page of the document. Upon return to Questar, the barcode information on each test was verified. Questar then followed up with the appropriate district(s) regarding any missing materials to ensure return or destruction (if materials were contaminated).

When the tests were delivered online, Test Examiners did not have access to the students' screens for the online assessment, only to the test administrator features. Students had unique, secure logins to access the MO EOC assessments they were registered for, and these logins were disabled after the student had tested. For tests with multiple sessions (those including a PE/WP), the students also had a Session Access code given to them by the teacher at the start of the session to ensure that students accessed the correct session of the test. Test items, as well as student responses, were encrypted during transmission to and from student computers.

3.7.2. Detection and Prevention of Testing Irregularities

To protect the validity and fairness of scores on the MO EOC assessments, DESE has implemented measures to prevent and detect cheating. Possible violations on the MO EOC assessments include the following:

- Copying and reviewing MO EOC assessment items with students
- Cueing students during testing either verbally or with written materials on the classroom walls
- Cueing students nonverbally, such as tapping or nodding the head
- Using a calculator on an EOC assessment that does not allow calculator use, unless specified by the student's IEP
- Using a calculator that contains stored equations or connects to the Internet
- Splitting sessions into two parts
- Ignoring the standardized directions in the test books
- Paraphrasing parts of the assessment to students

[5] **Standard 6.7:** Test users have the responsibility of protecting the security of test materials at all times (AERA, APA, NCME, 2014, p. 117).

- Changing or completing (or allowing other school personnel to change or complete) student answers
- Allowing accommodations that are not written in the IEP
- Allowing accommodations for students who do not have an IEP
- Allowing students to use dictionaries on parts of the MO EOC assessment other than the WP
- Defining terms on the test
- Allowing students to access cell phones or other electronic devices during testing

To detect cheating, DESE has implemented the following steps for the MO EOC assessments:

1. School officials, parents, and other interested parties call or e-mail DESE to report a testing concern or allegation.
2. A narrative of the conversation, if reported orally, is written and read back to the individual reporting the concern.
3. The superintendent of the district in which the allegation is made is then contacted and read the narrative or e-mail.
4. A letter is sent to confirm the conversation and to ask the superintendent to investigate the claim.
5. An MO EOC assessment Quality Assurance Concern District Response Report is sent for the superintendent to use for replying to the allegation.

DESE also implemented a self-monitoring process whereby District Test Coordinators completed a Quality Assurance (QA) self-monitoring form.[6] This QA process was issued to District Test Coordinators in an administrative memo.[7] The form was designed to be used by District Test Coordinators as part of their regular supervision process throughout the testing window, and it allowed districts to monitor and strengthen their administration of the MO EOC assessments. The questions on the form were designed to focus attention and help districts examine important areas of assessment training, administration, and test security.

District Test Coordinators were asked to complete one MO EOC Quality Assurance form for one EOC classroom. Regarding cheating prevention, the form asked District Test Coordinators to "Explain the district's test security plan" and answer the question, "What preventative measures are taken to curb cheating within the computer lab?" District Test Coordinators were urged to report testing irregularities or concerns immediately to the Assessment Section at assessment@dese.mo.gov or (573) 751-3545. DESE also performed onsite spot checks of quality assurance procedures during the spring testing window.

When testing irregularities were reported, DESE was able to request that Questar perform statistical analyses to detect and flag unusual response patterns. DESE then worked with districts to establish procedures for follow-up decisions appropriate to the situation.

---

[6] View the QA form online at http://tiny.cc/deseqaself2017.
[7] View the memo online at https://dese.mo.gov/sites/default/files/am/documents/CCR-17-001.pdf.

3.8. Test Administration

3.8.1. Test Organization

Students took the MO EOC assessments in one or two sessions depending on the content area. All assessments were administered online unless the student's IEP specified a Braille, Large Print, or Paper/Pencil administration. Each SR item consisted of a stem followed by four or more response options, and the student clicked an answer choice. The tests were not timed. Students were encouraged to complete an online tutorial of the online assessment platform prior to testing. This tutorial included instructions on how to use the tools in the system and practice questions for the students.

3.8.2. Test and Ancillary Materials

District Test Coordinators or School Test Coordinators were responsible for providing all MO EOC assessment materials to Test Examiners. The materials provided by Questar and/or DESE included the following:

- *Test Coordinator's Manual* (electronic copy)
- Large Print and/or Braille test materials
- Return kit materials for accommodated test materials
- Accommodated Paper/Pencil test booklets (printed from the online assessment platform by the school district)

Students taking an accommodated version of the MO EOC assessments needed the following additional materials, which were not provided by Questar or DESE:

- No. 2 pencils
- Scratch paper

For the online assessment, each student needed a computer with a monitor, mouse, and keyboard or a tablet device. Adequate space should have been left between workstations. Students could use scratch, grid, or draft paper and a writing utensil while taking the online assessment. The Test Examiner needed the following:

- A computer for logging on to the test administrator interface
- A writing board and utensil

Additionally, students taking either the Paper/Pencil or online version were allowed to use a calculator for the Algebra I, Algebra II, Geometry, Biology, and Physical Science assessments. Students taking the English I and English II writing prompts had access to a dictionary, thesaurus, and grammar handbook. Students taking any of the Mathematic assessments had access to the Mathematics Reference Sheet; students taking the Physical Science assessment had access to the Periodic Table of Elements. Starting in fall of 2018, students taking Biology had access to a Codon Wheel/Codon Table reference sheet.

Calculators could not contain stored equations or functions at the time of the EOC Mathematics and Science assessments. Test Examiners were responsible for ensuring and verifying that calculators with the ability to store functions and equations (e.g., a graphing or a scientific calculator) had the memory cleared before and after each Mathematics and Science assessment. Calculators could not have Internet connectivity or be able to connect to anyone inside or outside the classroom during testing. Students could not use a calculator on a laptop or other portable computer, pocket organizer, cell phone, device with a typewriter-style keyboard, electronic writing pad, or pen-input device unless a particular assistive device was required for a student and was specified on his or her IEP.

3.8.3. Preparing the Test Administration Site and the Students

Before students began the assessment using the online system, a representative of the district or school was responsible for the following tasks:

- Read the entire *Test Coordinator's Manual*.
- Review the DESE and Questar trainings regarding the EOC assessments.
- Run a workstation readiness test on each workstation used for testing.
- Ensure that the online test delivery system is downloaded to each workstation for test delivery.
- Provide an upload to DESE (precode file) of all students that will be testing for the current administration of the EOC assessments. (The precode file is a data file containing one record per student and each student is assigned a unique MOSIS ID. The purpose of the data file is to identify students, Examiners, and content areas for testing.)
- Input identification information for students who were not included in the precode file.
- Specify district testing windows within the Missouri statewide test administration window.

Additionally, the Test Examiner was responsible for setting and verifying class information and setting students' testing status codes and/or accommodations information in the online test administration system.

Students were NOT allowed to use electronic devices such as cellular phones, digital cameras, gaming devices, or scanners during the testing session. However, students could use calculators during the Algebra I, Algebra II, Geometry, Biology, and Physical Science test sessions. (See Section 3.7.2 for more information regarding calculator usage and restrictions.)

3.8.4. Directions for Administration

In accordance with Standard 6.1[8], specific standardized directions for administration were printed in the *Directions for Administration 2018–2019* (DFA) manual. Directions to be read aloud to the students were printed in bold type and had a callout arrow in the margin for clarity. Information for the teacher that should not be read aloud was in italic type. Figure 5.1 provides an example of a script from the DFA for the Geometry EOC assessment.

[8] **Standard 6.1:** Test administrators should follow carefully the standardized procedures for administration and scoring specified by the test developer and any instructions from the test user (AERA, APA, NCME, 2014; p. 114).

Geometry Directions–Session I

The Test Administrator should use his/her computer or tablet to create a testing session, as outlined in the *Test Coordinator's Manual*. On each student's device, open the secure browser to the student log-in page.

These directions pick up from where the "Preparing Students for Testing" section ended (page 13).

SAY **For the questions in this test, you will select an answer from the provided choices or type your answers in the space provided. Remember to check that you have correctly identified the answer choice after you select it. Your score on these questions will depend on how well you follow directions and show your understanding of Geometry. You may use a calculator and scratch paper to work through the questions. The Mathematics Reference Sheet is available during testing by selecting the REFERENCE SHEET icon. Select the HELP button for instructions on how to use the system tools.**

There are several important things to remember:

1. **Read each question carefully and think about the answer. Then choose the answer that you think is best.**
2. **If you do not know the answer to a question, mark it for review, skip it, and go on. You may return to it later.**
3. **When you finish the test, you may check your work.**

Do you have any questions?

Answer any questions the students may have.

SAY **You should now see a screen asking for the Access Code. In the space provided, type in** *[insert your Access Code]*. **Then select "Continue" to begin the test.**

If a student does not understand a word, you may pronounce the word for the student but do not define, explain, or paraphrase it. You may pronounce only one word per sentence. Pronouncing several words or phrases is an oral reading accommodation. If a student has not finished in the allotted time and is making adequate progress, the student should be allowed to finish. It is highly encouraged that all students complete the test during one testing session. However, should a student be unable to finish, testing may resume the following day. Test Examiners must address the concern with their STC immediately for assistance.

There is a separate Access Code for each session. Ensure the Access Code is used for the session being administered.

22

Figure 3.1. Directions for Administering from the DFA—Geometry

3.9. Accommodations and Modifications

A student's IEP team had the responsibility and authority to determine individual accommodations to support and ensure his or her participation in the MO EOC assessments. Students who were English language learners (ELLs) were also able to receive accommodations to support and ensure participation in the MO EOC assessments. The accommodations are intended to assist the student to demonstrate his or her knowledge, skills, and abilities. The accommodations available for the MO EOC assessments are listed in Appendix I. The accommodations for the MO EOC assessments include, but were not limited to, the following:

- A student may receive a modified version of the testing materials, such as the Braille, Large Print, or Paper/Pencil edition.
- A teacher may present the test content to a student in a nonstandard way, such as by reading it aloud in English or in the student's native language, paraphrasing it, or using sign language. For the English I and English II assessments, this will result in the lowest obtainable scale score (LOSS) being assigned.
- A student may be allowed additional time to complete one or more sessions of the assessment.
- A student may use an assistive communicative device.
- A student may be tested individually or in a small group.

- A student may be allowed to use a computer, another word-processing device, or a teacher scribe to record his or her responses.
- A student may use other assistive materials such as a bilingual dictionary.

Modifications are alterations in the test that change construct-related requirements. The resulting information may not be equal to the information that might be obtained without modifications. The following modifications for the MO EOC assessments were able to be provided:

- Oral reading of the assessment, including paraphrasing questions
- Oral reading in native language
- Use of a bilingual dictionary for the English I or English II assessment

In accordance with Standard 6.3[9], Test Examiners indicated an accommodation by checking the appropriate box(es) for the student in the online test administration site.

Tables 3.2 and 3.3 contain information about the percentage of students who received each type of accommodation for each MO EOC assessment for Fall 2018 and Spring 2019. The most prevalent type of accommodation for the Fall 2018 and Spring 2019 administrations across all MO EOC assessments was testing in "Other Setting." See Appendix I for a list of accommodation codes from the *2018–2019 Test Coordinator's Manual*.

[9] **Standard 6.3:** Changes or disruptions to standardized test administration procedures or scoring should be documented and reported to the test user (AERA, APA, NCME, 2014, p. 115).

Table 3.2. Accommodation Distributions—Fall 2018

Accommodation	English I		English II		Algebra I		Algebra II		Geometry		Biology		Physical Science	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Braille	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Large Print	0	0.00	0	0.00	1	0.02	0	0.00	0	0.00	1	0.04	0	0.00
Read Aloud (ELA Reading Passages) - Assistive	0	0.00	8	0.32	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Read Aloud (ELA Reading Passages) - Human Reader	0	0.00	47	1.85	16	0.30	0	0.00	0	0.00	9	0.36	0	0.00
Read Aloud (ELA Reading Passages) - Blind Students	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Signing of Assessment	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Paper Based Assessment	0	0.00	3	0.12	12	0.23	0	0.00	0	0.00	4	0.16	0	0.00
Read Aloud (ELA Reading Passages) - Native	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Scribe for Non-ELA Writing	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Speech-to-Text - Assistive Technology	0	0.00	1	0.04	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Abacus	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Multiplication Table	0	0.00	3	0.12	11	0.21	0	0.00	0	0.00	3	0.12	0	0.00
Specialized Calculator	0	0.00	1	0.04	6	0.11	0	0.00	0	0.00	0	0.00	0	0.00
Alternate Response Options	0	0.00	0	0.00	1	0.02	0	0.00	0	0.00	0	0.00	0	0.00
Read Aloud (Not Including ELA Reading Passages) –	10	6.13	114	4.49	249	4.69	3	0.55	1	0.72	172	6.88	0	0.00
Read Aloud (Not Including ELA Reading Passages) – Assistive Technology	0	0.00	7	0.28	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Read Aloud (Not Including ELA Reading Passages) –	0	0.00	41	1.62	66	1.24	3	0.55	0	0.00	22	0.88	0	0.00
Color Contrast - Paper	0	0.00	2	0.08	2	0.04	0	0.00	0	0.00	0	0.00	0	0.00
Color Overlay	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Magnification - Assistive Technology	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Masking - Paper	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Translation	0	0.00	9	0.35	6	0.11	0	0.00	0	0.00	3	0.12	0	0.00
Read Aloud (Not Including ELA Reading Passages) -	0	0.00	2	0.08	5	0.09	0	0.00	0	0.00	0	0.00	0	0.00
Scribe	0	0.00	5	0.20	3	0.06	0	0.00	0	0.00	5	0.20	0	0.00
Bilingual Dictionary	0	0.00	42	1.65	4	0.08	0	0.00	0	0.00	3	0.12	0	0.00
Separate Setting	15	9.20	129	5.08	269	5.07	4	0.73	1	0.72	142	5.68	0	0.00

Table 3.3. Accommodation Distributions—Spring 2019

Accommodation	English I		English II		Algebra I		Algebra II		Geometry		Biology		Physical Science	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Braille	0	0.00	3	0.00	2	0.00	1	0.01	1	0.03	6	0.01	0	0.00
Large Print	1	0.01	30	0.05	17	0.03	5	0.03	0	0.00	30	0.05	0	0.00
Read Aloud (ELA Reading Passages) - Assistive	0	0.00	95	0.15	11	0.02	0	0.00	0	0.00	24	0.04	0	0.00
Read Aloud (ELA Reading Passages) - Human Reader	349	3.14	1,634	2.65	334	0.56	6	0.04	7	0.20	417	0.69	37	1.62
Read Aloud (ELA Reading Passages) - Blind Students	1	0.01	1	0.00	2	0.00	1	0.01	0	0.00	0	0.00	0	0.00
Signing of Assessment	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Paper Based Assessment	1	0.01	99	0.16	100	0.17	2	0.01	1	0.03	68	0.11	0	0.00
Read Aloud (ELA Reading Passages) - Native Language	0	0.00	8	0.01	0	0.00	1	0.01	0	0.00	2	0.00	0	0.00
Scribe for Non-ELA Writing	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Speech-to-Text - Assistive Technology	1	0.01	59	0.10	68	0.11	0	0.00	1	0.03	56	0.09	0	0.00
Abacus	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Multiplication Table	0	0.00	15	0.02	187	0.31	1	0.01	1	0.03	48	0.08	2	0.09
Specialized Calculator	5	0.05	14	0.02	116	0.19	1	0.01	4	0.11	60	0.10	3	0.13
Alternate Response Options	0	0.00	5	0.01	5	0.01	0	0.00	0	0.00	6	0.01	0	0.00
Read Aloud (Not Including ELA Reading Passages) –	658	5.93	5,095	8.28	5,477	9.17	797	5.10	57	1.59	6,123	10.13	96	4.20
Read Aloud (Not Including ELA Reading Passages) – Assistive Technology	1	0.01	68	0.11	105	0.18	1	0.01	0	0.00	124	0.21	0	0.00
Read Aloud (Not Including ELA Reading Passages) –	174	1.57	861	1.40	802	1.34	25	0.16	44	1.23	1,111	1.84	28	1.22
Color Contrast - Paper	2	0.02	357	0.58	159	0.27	15	0.10	0	0.00	130	0.22	2	0.09
Color Overlay	58	0.52	10	0.02	2	0.00	0	0.00	1	0.03	3	0.00	1	0.04
Magnification - Assistive Technology	1	0.01	9	0.01	10	0.02	0	0.00	0	0.00	12	0.02	0	0.00
Masking - Paper	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
Translation	1	0.01	29	0.05	52	0.09	0	0.00	5	0.14	33	0.05	0	0.00
Read Aloud (Not Including ELA Reading Passages) -	1	0.01	27	0.04	43	0.07	2	0.01	5	0.14	33	0.05	0	0.00
Scribe	10	0.09	108	0.18	73	0.12	1	0.01	4	0.11	100	0.17	1	0.04
Bilingual Dictionary	17	0.15	66	0.11	28	0.05	0	0.00	0	0.00	37	0.06	0	0.00
Separate Setting	818	7.37	4,787	7.78	4,282	7.17	141	0.90	139	3.89	4,451	7.37	146	6.39

3.10. Materials Handling and Return

3.10.1. Materials Handling during Administration

The *Test Coordinator's Manual* contained detailed instructions for how schools and districts should collect and package the Paper/Pencil, Braille, and/or Large Print testing materials at the end of the test administration. For Test Examiners, these activities included, but were not limited to, the following:

- Collecting test books from the students using the accommodated editions
- Returning all used and unused test books to the School Test Coordinator
- Collecting all scratch paper used during testing
- Properly handling all contaminated test books (i.e., books having contact with bodily fluids such as blood or with any potentially hazardous material)

For School Test Coordinators, these activities included, but were not limited to, the following:

- Collecting testing materials from the Test Examiners
- Returning all test books (used and unused) to the District Test Coordinator
- Destroying all nonsecure testing materials

After receiving the used and unused test books from the School Test Coordinators, District Test Coordinators completed the following steps:

- Verifying 100% return of test books
- Completing the Test Book Accountability Form and faxing it to Questar

For the online system, the student needed to click the submit button once he or she had finished testing to submit the test for scoring. No additional information was needed from the Test Examiner after the student had completed the test. All demographic information was edited or added by the test administrator before the student started the assessment.

3.10.2. Questar's Secure Material Check-In Procedures

Questar adhered to strict quality assurance procedures in order to ensure that all accommodated test books were returned and accounted for. The check-in procedures included multiple steps to ensure that no test books were overlooked. All staff members received thorough and specific training before they participated in the check-in of test books.

Upon receipt of accommodated test books from the school districts, boxes were kept in a secure location and remained sealed until check-in. If a box had to be opened for any reason, it was immediately resealed.

Two teams checked in the secure materials. The first team prepared the test books for scanning. One district box was opened at a time, and secure test books were separated from ancillary materials and stacked on carts to be checked in. This process was repeated for all boxes for a district to ensure that all materials returned to Questar at the same time were checked in at the same time. Once the first team filled the cart(s) with all the secure materials from a district, the cart(s) was passed to a second team.

The second team checked in each test book by scanning the secure barcode into Questar's database. Operators worked in teams of two at computers equipped with barcode scanners. Operator 1 counted and scanned enough secure documents to fill a storage box. The operator verified that the database collected the same number of barcodes. If there was a discrepancy, an immediate reconciliation took place. Each ID number (barcode number) had a check digit that ensured that all numbers were correctly read by the scanner and that no ID number was miskeyed when manually entered. If a barcode was damaged or not readable, the operator manually entered the barcode number into the system. After this process was complete, the box of secure materials was handed to Operator 2 and scanned a second time. The database verified that the same barcode numbers were read during the scanning of the box or an immediate reconciliation took place. After verification, the secure materials were placed in a Questar box for storage. The scanning system provided audible and onscreen cues to alert operators of scanning discrepancies.

Further validity checks were done before each box was sealed to ensure that there were no ID barcode scanning discrepancies and that all ID numbers were correct. The validity checks also ensured that the ID numbers and the quantity in each box matched what was entered into the database. Finally, each box was placed on a pallet and stored.

Post-check-in procedures were also performed prior to notifying the districts of missing secure materials. For any district that was missing a secure material, an individual box-by-box hand search was conducted in an attempt to locate the secure material(s). If an unaccounted secure material was found, the material was then coded into the database by a Questar supervisor, and Questar's Program Management team was notified. If unaccounted-for material(s) were not found during the box-by-box hand search, the material(s) was considered missing and the district was notified via the Secure Missing Material Report process. This was also communicated to DESE, who would then follow up with discretion.

3.11. Summary

The distribution, administration, and collection of the MO EOC assessments were carefully communicated and executed in accordance with the detailed *Test Coordinator's Manual*. All standards related to test security, administration, and accommodations were adhered to throughout the process. The most important steps and procedures have been covered in this chapter. Readers interested in further detail should consult the *Test Coordinator's Manual* for the MO EOC assessments.

Chapter 4: Scoring

4.1. Introduction

The MO EOC assessment forms were processed and scored by Questar. SR items are automatically scored against a fixed key immediately after a test is submitted by the student. Each test form is tested entering 100% correct responses and 100% incorrect responses through both desktop and tablet clients, and each test score is validated as part of a comprehensive end-to-end process culminating in final reports.

The PE/WPs were scored by Questar's qualified scorers. This chapter outlines the processes used to implement scoring materials for the PE/WPs and CR, receive and scan student responses, hire and train scorers, score the PE/WPs and CRs, and maintain control of the quality of the scoring processes.

4.2. Scorer Training and Scoring Processes

Questar hand-scored items on the English I and English II assessments, which contained WPs, and items on the Algebra I, Algebra II, Geometry, Biology, and Physical Science assessments, which contained CRs. The PE/WPs required students to respond with extended written answers to questions on given topics or to questions regarding specific events.

The following sections Questar's processes for scoring the PE/WPs and CRs in the MO EOC assessments for 2018–2019, which were consistent for two administrations (Fall 2018 and Spring 2019). The PE/WPs and CRs were scored by human raters.

4.2.1. Scorer Training

4.2.1.1. <u>Recruitment and Selection</u>

Scoring quality starts with the recruitment process and extends through screening and placement (assigning scorers to prompts based on their skills and experience), training, qualification, and scoring. Questar accessed a large pool of educated candidates to professionally evaluate assessment prompts.

Questar selected scorers according to their strengths and background. All scorers had, at a minimum, a four-year college degree. The following steps show an overview of key processes:

1. <u>Process Timeline and Recruitment Tool:</u> Questar used a web-based application to collect data on scorer education, prior scoring experience, and other key information to screen candidates currently in the database system.
2. <u>Initial Screening:</u> Candidate data was analyzed, and prospective scorers prioritized.
3. <u>Offer:</u> Questar contacted prospective scorers detailing project requirements, timelines, and quality standards.
4. <u>Final Documentation and Project Placement:</u> Scorers signed confidentiality agreements to consent to keep all information and student responses confidential. Only scorers who successfully completed training and qualifying were allowed to evaluate student responses.

4.2.1.2. Training Materials

There was not a Summer 2018 administration. For the Fall 2018 and Spring 2019 administrations, Questar content specialists created training materials that were reviewed by DESE content specialists. Questar scoring staff communicated with DESE during this process regarding item questions or clarifications.

Training materials included the following:

- Anchor Sets: The anchor set is the primary reference for scorers as they internalize the rubric during training. All scorers had access to the anchor set while scoring and were directed to refer to it regularly.

- Practice Sets: Practice sets were used to help trainees develop experience in independently applying the scoring guide or rubric to student responses. The practice sets provided guidance and practice for trainees in defining the line between score points as well as applying the scoring criteria to a wider range of types of responses.

- Qualification Sets: All qualifying sets were used to confirm that scorer trainees had grasped the scoring criteria and were able to accurately assign the range of scores to student responses. Scorer trainees had to demonstrate acceptable performance on these sets by meeting a predetermined standard for accuracy to qualify to score MO EOC performance events and writing prompts. Questar's digital scoring system programmatically enforced qualification rules.

4.2.1.3. Training Process

Scorers went through training and qualifying prior to scoring on site, including reviewing scoring guidelines and procedures. This training provided scorers with a clear understanding of the training materials and scoring protocols of the MO EOC assessments. Scorers were expected to read and review annotations of the training materials with focused direction given by scoring directors or content specialists. The following are the steps used during the training of the items:

- Scoring for Questar: This gave a brief overview of what scoring is, the tools provided to help the scorers, and the individuals who would support the scorers during the project.

- Questar Scoring System: Scorers were trained on the internal scoring system.

- Scoring the Missouri Project: Specifics were provided regarding the Missouri Project. DESE and Questar worked collaboratively so the scorers understood the project.

- Scoring the Item: This training process walked the scorers through the anchor, practice, and qualification papers. The scorers proceeded through the qualification process and, upon qualifying, they continued on to operational scoring.

- Additional Training: Before operational scoring could begin, information on how to handle unscorable student responses and alert responses was provided.

Scoring started for the scorer once all of the steps were successfully completed.

4.2.1.4. Scorer Selection with Qualification

If applicants did not successfully complete the training and qualifying requirements, they were not allowed to score any MO EOC student responses. Furthermore, qualified scorers were dismissed if their scoring performance did not meet defined standards. Below are the qualification standards that must have been met in order to score the Missouri Project. The range of possible scores is noted below. The 4-point items have possible score points of 0, 1, 2, 3, and 4; or 1, 2, 3, and 4, depending on the item. All other hand-scored items have 0 as the lowest score. Exact agreement is based on agreement to the answer key. Depending on the prompt, there will be either one or two qualifying sets. For prompts that require a specific response or responses, as may be found in math, one qualifying set is used.

- 4-point items
    - (0–4, 1–4)
    - 2 sets of 10 papers
    - 80% exact agreement on one of two sets
    - Scorers saw both sets. If they passed the first, the second was a review.

- 3-point items
    - (0–3)
    - 1 or 2 sets of 10 papers, depending on the item.
    - 80% exact agreement on one set.
    - If two sets, scorers saw both sets. If they passed the first, the second was a review.

- 2-point items
    - (0–2)
    - 1 or 2 sets of 10 papers, depending on the item.
    - 80% exact agreement on one set.
    - If two sets, scorers saw both sets. If they passed the first, the second was a review.

- 1-point items
    - (0–1)
    - 1 or 2 sets of 10 papers, depending on the item.
    - 80% exact agreement on one set.
    - If two sets, scorers saw both sets. If they passed the first, the second was a review.

4.2.1.5. Second Read Procedures

Rater agreement is the agreement between the first and second scores assigned to student responses. Rater agreement indices include exact, adjacent, and nonadjacent agreement. Guidelines for rater agreement are determined in accordance with customer requirements and Questar scoring standards for exact and adjacent agreement. Questar scoring staff used rater agreement results as one factor in determining the needs for continuing training and intervention on individual levels.

Questar's scoring system included comprehensive rater agreement reports that allowed scoring directors to monitor both individual and group performance. Responses were randomly assigned to scorer one. After the first score was applied, the system automatically sent the tenth document to a different scorer for a second read. Scorer one provided the score of record, and the second read was for rater agreement purposes only.

4.2.2. Scoring Processes with Monitoring and Recalibration Procedures

4.2.2.1. Read-Behinds

The process of reading behind scorers (hereafter referred to as a read-behind) was a major responsibility of Questar's content staff and a primary tool for guarding against scorer drift. Questar's scoring system's integrated read-behind tool allowed Questar staff to review the scores assigned to individual student responses by any given scorer. The team leads used an internal report to monitor and ensure consistent scoring. If an incorrect score was identified during the read-behind, the correct score was assigned; that score became the score of record. Questar staff performed random read-behinds on every scorer through-out the scoring window, as well as targeted read-behinds based on scorer monitoring reports. Questar staff perform an average of 3-5% read-behinds on every scorer.

Questar's content staff could perform a search for the following:

- Responses scored by a particular scorer
- Responses receiving a particular score point
- Responses with scores that agree with, are adjacent to, or are nonadjacent to each other
- Combinations of these features

Content staff reviewed responses to confirm that the scores were correctly assigned and to give customized feedback and remediation to individual scorers.

4.2.2.2. Calibration

Content staff used calibration sets as needed to reinforce scoring standards, introduce scoring decisions, or correct scoring issues and trends. The primary goal of calibration was to continue training and to reinforce the scoring standards. Calibration sets may be "on the line" between score points or might contain unusual examples that are challenging to score and therefore useful for reinforcing the scoring rubric. Online calibration sets could be given to entire groups, a subset of scorers, or individual scorers, as needed, to score independently. These annotated sample responses promoted accuracy by exploring project-specific issues, score boundaries, or types of responses that were particularly challenging to score consistently. After scoring an online calibration set, scorers could ask questions and seek clarification of the score point or annotation.

Calibration sets are developed throughout the scoring window, using responses that serve as training examples both in one-on-one and group situations. Calibration sets are also used after a weekend off, if needed. These papers are shredded after the project is complete.

4.2.2.3. Managing Scoring Quality (Scorer Exception Processing)

Content staff, often along with a project manager or human resource representative, intervened when scorer performance statistics did not meet quality standards, or a scorer violated other Questar policies. Intervention included calibration, retraining, direct counseling and review of papers, and requalification. Scorer exception processing allowed Questar's project managers to define intervals at which the scoring system would check scorer validity for exact and adjacent agreement. If scorers were below pre-set standards, staff monitoring this process would interrupt their scoring process to review anchor papers or take other steps to improve their scoring. Through this process, Questar's scoring system could provide an additional training/requalification set and, if performance was not improved, could lock scorers out of the scoring system. This process prevented scorers from continuing to score if standards were not maintained.

Because the system monitored scorers and provided the scorers' information quickly, Questar's content staff continually focused on quality-control measures. These measures included read-behinds, calibration, and responding to questions in the review queue. Content staff were able to spend more time working directly with scorers who had questions.

4.3. Quality Control with Validity Responses

Validity responses are pre-scored responses strategically interspersed in the pool of operational responses. These responses are not distinguishable from operational responses, and scorers' scores are only accepted for monitoring purposes, not in replacement of the true score.

The use of validity responses provides an objective procedure that helps ensure that scorers are applying the same standards throughout the project. This procedure offers feedback on the accuracy and consistency of individual scorers and groups of scorers assigned to a given item. Questar's validity mechanism provides an objective and systematic check of accuracy. It verifies that scorers are applying the same standards throughout the project and, therefore, guards against scorer drift and ultimately group drift. This procedure provides immediate feedback on individual scorers and the group as a whole.

Validity papers are actual student responses chosen by the team leads or scoring director as examples that clearly earn certain scores. There is only one scoring director per content area. Following the standards established, scoring directors assigned "true scores" to validity responses to compare how often scorers match them throughout the scoring session. The validity pool included responses encompassing the entire score range for each item. Scorers scored them without being aware they were scoring validity papers rather than operational responses. Validity responses were sent to scorers throughout the project. They were selected and corroborated by scoring leaders.

Each MO EOC content area was set to contain validity papers at a frequency rate determined by the range of scores and complexity of each item. This means that each scorer would see a validity paper at varying times throughout the project. The scorers could not distinguish a validity paper from an operational response because these papers are pulled from operational scoring. The process of selecting validity papers and refreshing the pool was to select papers scored by expert readers. Questar's system allows a team leader, scoring director, or content specialist to score validity items using a hierarchical approval process to ensure the score has been adequately confirmed. For instance, if a score of 3 was given by a team leader, it could not be selected for a validity response unless confirmed and approved by the scoring director. If the validity response was chosen by the scoring director, the response must be confirmed and approved by the content specialist.

Tables 4.1 and 4.2 show the summaries of the validity paper results at the end of the project for the Fall 2018 and Spring 2019 administrations, respectively. The “Rater Agreement Plan” column indicates the expected percentage of agreement given the maximum points available for the item. A higher percentage was expected for items with fewer points, and a lower percentage was expected for items with more points. For example, the rater agreement plan was 100 percent for 1-point items and 80 percent for 4-point items. The “Rater Agreement Actual” column shows the observed rater agreement. The variance is the difference between the actual and planned rater agreement. Positive values indicate that the actual agreement was higher than the planned agreement, whereas negative values indicate that the actual agreement was lower than the planned agreement. Overall, items worth 1 point tended to have smaller variances than items worth 2 or more points. The item with the largest negative variance in the Fall administration was an item where rater agreement was lower than the planned rater agreement (88% versus 95% for a 2-point English I item).The item with the largest negative variance in the spring was an item where rater agreement was lower than the planned rater agreement (82% versus 100% for a 1-point Physical Science item). There are many incidents of positive variance, where the rater agreement was higher than the planned rater agreement. The results of the validity paper scoring indicate that for the majority of items, the variance was relatively small and within 10%.

Table 4.1. Summary of Validity Paper Results—Fall 2018

Item	*n*-Count Responses Received	*n*-Count Responses Scored	#Points	Rater Agreement Plan	Rater Agreement Actual	Variance
English I - MO0001881-1	172	167	4	80%	75%	-5%
English I - MO0001881-2	172	167	4	80%	88%	+8%
English I - MO0001881-3	172	167	2	95%	88%	-7%
English II - MO0001805-1	2,524	2,431	4	80%	82%	+2
English II - MO0001805-2	2,524	2,431	4	80%	79%	-1%
English II - MO0001805-3	2,524	2,431	2	95%	93%	-2%
Algebra I - MOA116353	5,311	1722	1	100%	100%	0%
Algebra I - MOA116493	4,556	1728	1	100%	100%	0%
Algebra I - MOA116502	3,924	532	1	100%	100%	0%
Algebra I - MOA116581_1-1	4,523	2,488	1	100%	100%	0%
Algebra I - MOA116581_1-2	4,523	2,488	1	100%	100%	0%
Algebra I - MOA116581_1-3	4,523	2,488	1	100%	100%	0%
Algebra I - MOA116581_6	5,290	4847	3	85%	96%	+11%
Algebra II - MO0001073	285	284	1	100%	100%	0%
Algebra II - MO0001083-1	556	547	2	95%	98%	+3%
Algebra II - MO0001083-2	556	547	2	95%	100%	+5%
Algebra II - MOA2163	455	385	1	100%	100%	0%

Item	*n*-Count Responses Received	*n*-Count Responses Scored	#Points	Rater Agreement Plan	Rater Agreement Actual	Variance
Algebra II - MOA21673	249	209	1	100%	100%	0%
Algebra II - MOA21695	517	61	1	100%	100%	0%
Geometry - MO0001747-1	70	70	1	100%	100%	0%
Geometry - MO0001747-2	70	70	1	100%	100%	0%
Geometry - MO0033067-1	117	117	1	100%	100%	0%
Geometry - MO0033067-2	117	117	1	100%	100%	0%
Geometry - MOG1624	118	117	1	100%	100%	0%
Geometry - MOG16224	110	109	1	100%	100%	0%
Geometry - MOG16294	49	48	1	100%	100%	0%
Geometry - MOG16508	50	49	1	100%	100%	0%
Geometry - MOG16576	94	93	1	100%	100%	0%
Geometry - MOG16777	81	80	1	100%	100%	0%
Geometry - MOG16791	59	58	1	100%	100%	0%
Geometry - MOG16812	106	105	1	100%	100%	0%
Biology - MO0008484	800	316	1	100%	100%	0%
Biology - MO0008963	1,191	1,119	1	100%	97%	-3%
Biology - MO0030876	1,193	1,138	1	100%	97%	-3%
Physical Science - MO0007638	50	50	1	100%	100%	0%
Physical Science - MO0007638	50	50	1	100%	100%	0%
Physical Science - MO0007638	28	28	1	100%	100%	0%

Table 4.2. Summary of Validity Paper Results—Spring 2019

Item	*n*-Count Responses Received	*n*-Count Responses Scored	#Points	Rater Agreement Plan	Rater Agreement Actual	Variance
English I - MO0001870-1	5,631	5,555	4	80%	73%	-7%
English I - MO0001870-2	5,631	5,555	4	80%	72%	-8%
English I - MO0001870-3	5,631	5,555	2	95%	93%	-2%
English I - MO0001873-1	6,318	6,238	4	80%	75%	-5%
English I - MO0001873-2	6,318	6,238	4	80%	72%	-8%
English I - MO0001873-3	6,318	6,238	2	95%	88%	-7%
English II - MO0001817-1	27,686	27,525	4	80%	75%	-5%
English II - MO0001817-2	27,686	27,525	4	80%	70%	-10%
English II - MO0001817-3	27,686	27,525	2	95%	92%	-3%
English II - MO0001843-1	32,983	32,861	4	80%	82%	+2%
English II - MO0001843-2	32,983	32,861	4	80%	73%	-7%
English II - MO0001843-3	32,983	32,861	2	95%	92%	-3%
Algebra I - MO0007286	59,120	26,739	1	100%	99%	-1%
Algebra I - MO0007870	32,387	31,751	1	100%	96%	-4%
Algebra I - MO0008294	26,734	26,359	1	100%	97%	-3%

Item	*n*-Count Responses Received	*n*-Count Responses Scored	#Points	Rater Agreement Plan	Rater Agreement Actual	Variance
Algebra I - MO0008357	26,735	18,259	1	100%	98%	-2%
Algebra I - MO0008437	32,387	12,085	1	100%	96%	-4%
Algebra I - MO0008732-1	32,191	31,402	2	95%	83%	-12%
Algebra I - MO0008732-2	32,191	31,402	2	95%	84%	-11%
Algebra I - MO0008734-1	32,191	17,071	1	100%	97%	-3%
Algebra I - MO0008734-2	32,191	17,071	1	100%	99%	-1%
Algebra I - MO0008762	26,577	25,360	3	85%	87%	+2%
Algebra I - MO0008925	26,634	3,509	1	100%	99%	-1%
Algebra I - MO0008998	32,386	31,551	1	100%	99%	-1%
Algebra II - MO0007277	8,119	1,685	1	100%	100%	0%
Algebra II - MO0007843	15,660	5,238	1	100%	99%	-1%
Algebra II - MO0007868	7,541	1,878	1	100%	100%	0%
Algebra II - MO0008067-1	8,097	7,933	1	100%	98%	-2%
Algebra II - MO0008067-2	8,097	7,933	1	100%	96%	-4%
Algebra II - MO0008216	8,119	2,037	1	100%	99%	-1%
Algebra II - MO0008281-1	7,523	6,845	1	100%	100%	0%
Algebra II - MO0008281-2	7,523	6,845	1	100%	99%	-1%
Algebra II - MO0008298	7,523	2,474	1	100%	99%	-1%
Algebra II - MO0008314	8,096	1,940	1	100%	99%	-1%
Algebra II - MO0020625-1	8,119	4,031	1	100%	100%	0%
Algebra II - MO0020625-2	8,119	4,031	1	100%	100%	0%
Algebra II - MO0020625-3	8,119	4,031	1	100%	100%	0%
Algebra II - MOA216353	8,119	1,334	1	100%	100%	0%
Geometry - MO0007943	1,738	1,718	2	95%	99%	+4%
Geometry - MO0007978	1,859	1,623	3	85%	99%	+14%
Geometry - MO0008022	1,859	1,798	3	85%	98%	+13%
Geometry - MO0008029	1,868	1,031	1	100%	100%	0%
Geometry - MO0008046	1,738	885	1	100%	100%	0%
Geometry - MO0008829	1,872	712	1	100%	100%	0%
Geometry - MO0008829	1,872	712	1	100%	100%	0%
Geometry - MO0008861	1,738	760	1	100%	100%	0%
Geometry - MO0008946	1,736	409	1	100%	100%	0%
Biology - MO0008484	26,607	5,173	1	100%	97%	-3%
Biology - MO0008963	26,607	20,869	1	100%	89%	-11%
Biology - MO0030876	26,607	26,253	1	100%	96%	-4%
Physical Science - MO0007638	2,482	2,327	1	100%	87%	-13%
Physical Science - MO0008212	2,482	2,350	1	100%	82%	-18%
Physical Science - MO0008421	2,482	737	1	100%	96%	-4%

4.3.1. Validity as Review

Selected validity responses were annotated by the content staff and flagged for review. If a scorer incorrectly scored one of these responses, content staff would address this with the scorer. This feedback helped to prevent scorer drift. Once a scorer received a validity response, it was not re-administered to that scorer.

4.3.2. Frequency Distribution

Frequency distribution, or the number or percentage of scores assigned at each score point of a rubric, was another key metric tracked and managed during scoring. Questar evaluated any anomalous scoring trends at the item and scorer level and intervened with the individuals involved. Anomalous scoring trends were determined by comparing individual reader distribution of scores to the overall group distribution of scores. Frequency distribution reports showed a breakdown of score points assigned on a given item. Expressed in percentages, data in these reports showed how often scorers, individually and as a group, assigned each score point.

4.3.3. Retraining and Resetting Scores

Questar's electronic scoring system has the ability to purge the scores assigned by a scorer whose work was deemed substandard. In those cases, the scores assigned by that individual would be cleared from the database. The responses would then be rerouted to qualified scorers and rescored according to the original scoring design. The scoring system also allows scoring leadership to reset scores for a date range or an item. Questar has not had to use this process to-date during this project. If it had, that reader would have been removed from the project.

4.3.4. Reporting and Data Analysis

Questar's digital scoring system automatically captured and tracked all score data. By reviewing up-to-date scorer performance statistics, Questar could quickly identify particular scorers whose performance fell outside of group norms while also keeping close track of the group as a whole. Reports for use in quality monitoring and project completion status were generated and updated automatically and were available to Questar scoring leadership staff at any time via the digital scoring system. Questar's reports gave daily and cumulative statistics and provided individual and group average agreement percentages.

4.3.5. Item Types and Score Points for Each Content Area

4.3.5.1. Fall 2018

English I and English II contained blended essay prompts of argumentative, expository, and/or narrative genres, with score points of 1–4, 1–4, and 0–2, respectively.

Algebra I, Algebra II, and Geometry contained constructed responses of 0–1, 0–2, 0–3, and 0–4 score points.

Biology and Physical Science contained constructed responses of 0–1 score points.

4.3.5.2. Spring 2019

English I and English II contained blended essay prompts of argumentative, expository, and/or narrative genres, with score points of 1–4, 1–4, and 0–2, respectively.

Algebra I, Algebra II, and Geometry contained constructed responses of 0–1, 0–2, 0–3, and 0–4 score points.

Biology and Physical Science contained constructed responses of 0–1 score points.

4.4. Quality Measure of Scoring: Rater Agreement

Rater agreement provides evidence supporting scorer consistency. Tables 4.4 and 4.5 present the rater agreement for each item for Fall 2018 and Spring 2019, respectively. The tables provide the total *n*-count for each item; the *n*-count for each item minus the number of auto-scores, expert scores, and read-behinds; and the *n*-count of double reads (i.e., the responses that received a second read). The actual agreement rates are raw data-rates before any adjudication was performed, if needed, by team leads and scoring directors and were calculated based on the double reads. The percentage of student responses of which two raters agreed exactly for a given item is presented (Exact Agreement Actual). Some degree of disagreement is to be expected with human judges, so the Exact + Adjacent Agreement is presented. For a few 1-point Algebra I, Algebra II, and Geometry items, the Exact + Adjacent Agreement is less than 100% due to instances where the raters disagreed on whether the student response was scoreable.

Across the two administrations, the Exact Agreements were higher than 80% with a few exceptions, and the Exact + Adjacent Agreements were perfect (100%) for the majority of items and 95% or higher for all but one of the remaining items.

Another approach to rater agreement is weighted kappa, which corrects for chance agreement (i.e., the probability that two raters will agree simply by chance based on number of score points available). The Fleiss-Cohen weights were applied for the weighted kappa statistic (Fleiss & Cohen, 1973). Landis and Koch (1977) proposed the following interpretation guidelines for kappa values (see Table 4.3).

Table 4.3. Guidelines for Weighted Kappa Values

κ	Interpretation
<0	Poor agreement
0.01 to 0.20	Slight agreement
0.21 to 0.40	Fair agreement
0.41 to 0.60	Moderate agreement
0.61 to 0.80	Substantial agreement
0.81 to 1.00	Almost perfect agreement

Across the two administrations, most of the weighted kappa values fall within the category of Almost Perfect Agreement (49 of 59, 83%), although there are a few items in the Substantial (6 of 59, 10%) and Moderate (4 of 59, 7%) Agreement classifications. In summary, the rater agreement percentages (i.e., Exact, Exact + Adjacent) and the weighted kappa results indicate a high degree of consensus among raters for the hand-scored items.

Table 4.4. Rater Agreement—Fall 2018

Item	# of Score Points	*n*-Count Responses Received	*n*-Count Responses Scored	*n*-Count of Double Reads	Exact Agreement Plan	Exact Agreement Actual	Exact + Adjacent Plan	Exact + Adjacent Actual	Weighted Kappa
English I - MO0001881-1	4	172	167	17	80%	75%	100%	100%	0.77
English I - MO0001881-2	4	172	167	17	80%	88%	100%	100%	0.89
English I - MO0001881-3	2	172	167	17	95%	88%	100%	100%	0.68
English II - MO0001805-1	4	2,524	2,431	228	80%	82%	100%	99%	0.86
English II - MO0001805-2	4	2,524	2,431	228	80%	79%	100%	99%	0.86
English II - MO0001805-3	2	2,524	2,431	228	95%	93%	100%	99%	0.85
Algebra I - MOA116353	1	5,311	1722	145	100%	100%	100%	100%	1.00
Algebra I - MOA116493	1	4,556	1728	152	100%	100%	100%	100%	1.00
Algebra I - MOA116502	1	3,924	532	36	100%	100%	100%	100%	1.00
Algebra I - MOA116581_1-1	1	4,523	2,488	221	100%	100%	100%	100%	1.00
Algebra I - MOA116581_1-2	1	4,523	2,488	221	100%	100%	100%	100%	1.00
Algebra I - MOA116581_1-3	1	4,523	2,488	221	100%	100%	100%	100%	1.00
Algebra I - MOA116581_6	3	5,290	4847	449	85%	96%	100%	100%	0.98
Algebra II - MO0001073	1	285	284	28	100%	100%	100%	100%	1.00
Algebra II - MO0001083-1	2	556	547	52	95%	98%	100%	100%	0.96
Algebra II - MO0001083-2	2	556	547	52	95%	100%	100%	100%	1.00
Algebra II - MOA2163	1	455	385	36	100%	100%	100%	100%	1.00
Algebra II - MOA21673	1	249	209	21	100%	100%	100%	100%	1.00
Algebra II - MOA21695	1	517	61	5	100%	100%	100%	100%	1.00
Geometry - MO0001747-1	1	70	70	4	100%	100%	100%	100%	1.00
Geometry - MO0001747-2	1	70	70	4	100%	100%	100%	100%	1.00
Geometry - MO0033067-1	1	117	117	8	100%	100%	100%	100%	1.00
Geometry - MO0033067-2	1	117	117	8	100%	100%	100%	100%	1.00
Geometry - MOG1624	1	118	117	7	100%	100%	100%	100%	1.00
Geometry - MOG16224	1	110	109	6	100%	100%	100%	100%	1.00
Geometry - MOG16294	1	49	48	3	100%	100%	100%	100%	1.00

Item	# of Score Points	*n*-Count Responses Received	*n*-Count Responses Scored	*n*-Count of Double Reads	Exact Agreement Plan	Exact Agreement Actual	Exact + Adjacent Plan	Exact + Adjacent Actual	Weighted Kappa
Geometry - MOG16508	1	50	49	3	100%	100%	100%	100%	1.00
Geometry - MOG16576	1	94	93	7	100%	100%	100%	100%	1.00
Geometry - MOG16777	1	81	80	6	100%	100%	100%	100%	1.00
Geometry - MOG16791	1	59	58	4	100%	100%	100%	100%	1.00
Geometry - MOG16812	1	106	105	9	100%	100%	100%	100%	1.00
Biology - MO0008484	1	800	316	29	100%	100%	100%	100%	1.00
Biology - MO0008963	1	1,191	1,119	113	100%	97%	100%	100%	1.00
Biology - MO0030876	1	1,193	1,138	104	100%	97%	100%	100%	1.00
Physical Science - MO0007638	1	50	50	3	100%	100%	100%	100%	1.00
Physical Science - MO0007638	1	50	50	3	100%	100%	100%	100%	1.00
Physical Science - MO0007638	1	28	28	1	100%	100%	100%	100%	1.00

Table 4.5. Rater Agreement—Spring 2019

Item	# of Score Points	*n*-Count Responses Received	*n*-Count Responses Scored	*n*-Count of Double Reads	Exact Agreement Plan	Exact Agreement Actual	Exact + Adjacent Plan	Exact + Adjacent Actual	Weighted Kappa
English I - MO0001870-1	4	5,631	5,555	499	80%	73%	100%	99%	0.68
English I - MO0001870-2	4	5,631	5,555	499	80%	72%	100%	99%	0.68
English I - MO0001870-3	2	5,631	5,555	499	95%	93%	100%	99%	0.58
English I - MO0001873-1	4	6,318	6,238	554	80%	75%	100%	99%	0.79
English I - MO0001873-2	4	6,318	6,238	554	80%	72%	100%	99%	0.79
English I - MO0001873-3	2	6,318	6,238	554	95%	88%	100%	99%	0.60
English II - MO0001817-1	4	27,686	27,525	2653	80%	75%	100%	98%	0.80
English II - MO0001817-2	4	27,686	27,525	2653	80%	70%	100%	98%	0.76
English II - MO0001817-3	2	27,686	27,525	2653	95%	92%	100%	99%	0.64
English II - MO0001843-1	4	32,983	32,861	3193	80%	82%	100%	99%	0.85
English II - MO0001843-2	4	32,983	32,861	3193	80%	73%	100%	99%	0.77

Item	# of Score Points	*n*-Count Responses Received	*n*-Count Responses Scored	*n*-Count of Double Reads	Exact Agreement Plan	Exact Agreement Actual	Exact + Adjacent Plan	Exact + Adjacent Actual	Weighted Kappa
English II - MO0001843-3	2	32,983	32,861	3193	95%	92%	100%	100%	0.65
Algebra I - MO0007286	1	59,120	26,739	1,623	100%	99%	100%	99%	1.00
Algebra I - MO0007870	1	32,387	31,751	3,129	100%	96%	100%	100%	1.00
Algebra I - MO0008294	1	26,734	26,359	2,601	100%	97%	100%	100%	1.00
Algebra I - MO0008357	1	26,735	18,259	1,355	100%	98%	100%	99%	1.00
Algebra I - MO0008437	1	32,387	12,085	692	100%	96%	100%	100%	1.00
Algebra I - MO0008732-1	2	32,191	31,402	3,058	95%	83%	100%	99%	0.89
Algebra I - MO0008732-2	2	32,191	31,402	3,058	95%	84%	100%	99%	0.90
Algebra I - MO0008734-1	1	32,191	17,071	1,047	100%	97%	100%	100%	1.00
Algebra I - MO0008734-2	1	32,191	17,071	1,047	100%	99%	100%	100%	1.00
Algebra I - MO0008762	3	26,577	25,360	2,430	85%	87%	100%	98%	0.92
Algebra I - MO0008925	1	26,634	3,509	259	100%	99%	100%	100%	1.00
Algebra I - MO0008998	1	32,386	31,551	3,112	100%	99%	100%	100%	1.00
Algebra II - MO0007277	1	8,119	1,685	137	100%	100%	100%	100%	1.00
Algebra II - MO0007843	1	15,660	5,238	347	100%	99%	100%	100%	1.00
Algebra II - MO0007868	1	7,541	1,878	132	100%	100%	100%	100%	1.00
Algebra II - MO0008067-1	1	8,097	7,933	735	100%	98%	100%	100%	1.00
Algebra II - MO0008067-2	1	8,097	7,933	735	100%	96%	100%	100%	1.00
Algebra II - MO0008216	1	8,119	2,037	134	100%	99%	100%	100%	1.00
Algebra II - MO0008281-1	1	7,523	6,845	645	100%	100%	100%	100%	1.00
Algebra II - MO0008281-2	1	7,523	6,845	645	100%	99%	100%	99%	1.00
Algebra II - MO0008298	1	7,523	2,474	155	100%	99%	100%	99%	1.00
Algebra II - MO0008314	1	8,096	1,940	120	100%	99%	100%	100%	1.00
Algebra II - MO0020625-1	1	8,119	4,031	279	100%	100%	100%	100%	1.00
Algebra II - MO0020625-2	1	8,119	4,031	279	100%	100%	100%	100%	1.00
Algebra II - MO0020625-3	1	8,119	4,031	279	100%	100%	100%	100%	1.00
Algebra II - MOA216353	1	8,119	1,334	114	100%	100%	100%	100%	1.00

Item	# of Score Points	*n*-Count Responses Received	*n*-Count Responses Scored	*n*-Count of Double Reads	Exact Agreement Plan	Exact Agreement Actual	Exact + Adjacent Plan	Exact + Adjacent Actual	Weighted Kappa
Geometry - MO0007943	2	1,738	1,718	156	95%	99%	100%	100%	1.00
Geometry - MO0007978	3	1,859	1,623	134	85%	99%	100%	100%	1.00
Geometry - MO0008022	3	1,859	1,798	170	85%	98%	100%	100%	0.99
Geometry - MO0008029	1	1,868	1,031	61	100%	100%	100%	100%	1.00
Geometry - MO0008046	1	1,738	885	48	100%	100%	100%	100%	1.00
Geometry - MO0008829	1	1,872	712	38	100%	100%	100%	100%	1.00
Geometry - MO0008829	1	1,872	712	38	100%	100%	100%	100%	1.00
Geometry - MO0008861	1	1,738	760	41	100%	100%	100%	100%	1.00
Geometry - MO0008946	1	1,736	409	26	100%	100%	100%	100%	1.00
Biology - MO0008484	1	26,607	5,173	366	100%	97%	100%	99%	1.00
Biology - MO0008963	1	26,607	20,869	1,777	100%	89%	100%	100%	1.00
Biology - MO0030876	1	26,607	26,253	2,588	100%	96%	100%	99%	1.00
Physical Science - MO0007638	1	2,482	2,327	230	100%	87%	100%	99%	1.00
Physical Science - MO0008212	1	2,482	2,350	233	100%	83%	100%	99%	1.00
Physical Science - MO0008421	1	2,482	737	54	100%	96%	100%	96%	1.00

Chapter 5: Psychometric Analyses

5.1. Overview of the Operational Test Data Analysis

Psychometric analyses are a pivotal part of the validation of the MO EOC assessments. This chapter provides the classical item statistics, the differential item functioning analysis results, IRT based scaling and equating procedures, and the information to evaluate the equivalency among test forms.

5.2. Classical Item Statistics

Classical item analysis is the analysis of item-level statistical information based on the classical data analysis. It is important to verify that the items and test forms function as intended. If any serious error were to occur with an item, errors should be flagged and evaluated for rectification (suppression, credit, or other acceptable solution) during item analysis. As a part of item level statistical analysis, classical differential item functioning (DIF) analysis is conducted at this stage. DIF analysis includes computation of standardized mean differences and Mantel-Haenszel statistics for MO EOC items to identify potential item bias. The item-level statistical analyses are based on the operational administrations that included responses from 23,799 students for Fall 2018, and 271,125 students for Spring 2019 across all content areas, as shown in Table 5.1. All classical data analysis results contribute information on the validity and reliability of the tests in Chapter 7.

Table 5.1. Number of Students Included in the Analyses

Test Period	Content Area	*n*-Count
Fall 2018	English I	163
	English II	2,538
	Algebra I	5,307
	Algebra II	545
	Geometry	138
	Biology	2,501
	Physical Science	39
	Government	12,514
	Am. History	54
	Total	23,799
Spring 2019	English I	11,101
	English II	61,555
	Algebra I	59,757
	Algebra II	15,636
	Geometry	3,576
	Biology	60,421
	Physical Science	2,286
	Government	51,987
	Am. History	4,806
	Total	271,125

5.2.1. Item-Level Statistics

Item difficulty is the proportion of students who gave correct responses to the item (also referred to as *p*-value) for dichotomous items. For polytomous items, the mean score is the average of the scores for students who responded to these items. The discrimination index is the point-biserial correlation between the item score and the total score based on the remaining items (also referred to as corrected point-biserial correlation). The student counts given are the total test population for that content area.

When building a test form for the MO EOC assessment, care is taken to refrain from choosing items with *p*-values less than 0.35 or greater than 0.95, or with negative point biserials. When *p*-values and point biserials are out of range, the answer keys are checked to verify that they are correct.

Appendix D presents the item difficulty, discrimination, and omit rates for all operational items on each assessment for the Fall 2018 and Spring 2019 administrations. Field test items are not included in the tables. The results indicate that the items measure achievement across a range of difficulty and that most items are correlated with the total test score, thereby discriminating between low- and high-performing students.

5.2.2. Speededness and Omit Rates

The consequence of time limits on students' scores is called *speededness*. A test is speeded if examinees do not have time to respond to all items on the test. Examinees may receive a lower score than they would have had the test not been timed. Most *speededness* statistics are based on the number of items that were not attempted by students. The MO EOC assessments were not designed to be speeded tests. Rather, they were intended to be *power tests*; that is, students are expected to have ample time to finish all items and prompts. For the purpose of this analysis, if a student did not attempt the last item on any of the separately timed subsections of the test, it was assumed that the student might not have reached the item because he or she ran out of time.

Item omit rates, especially for items appearing later in a test, are a gauge of potential test speededness. The "Omit Rate" column in Appendix D tables shows the percentage of students who omitted each SR item for each MO EOC assessment. As shown in the tables, the omit rates are zero or negligible for most items, thereby supporting the interpretation that the MO EOC assessments are power tests.

5.2.3. Differential Item Functioning (DIF)

Differential item functioning (DIF) occurs when an item has difficulty measures that vary substantially across subgroups of examinees with comparable ability. DIF will be examined using the Mantel-Haenszel (MH) (1959) procedure for dichotomous items and standardized mean differences for the polytomous items.

The Mantel-Haenszel method is a nonparametric approach to DIF. In the MH procedure, total raw scores are held constant while an odds ratio is estimated. In practice, the odds ratio is generally converted to the delta metric, and the Educational Testing Service (ETS) categorization is applied to flag the significance of DIF effects (Dorans & Holland, 1992).

With the groups matched on raw score, comparable examinees can be placed in j 2 × 2 tables of group by item response, where j equals the number of levels of the matching variable. For these analyses, if j equals each observed score category of the k-item tests, with $j = 0, 1, 2\ldots, k$, then one 2 × 2 table for a given item with score category j can be represented as the following:

Table 5.2. General Notation for the 2 x 2 Data Matrix

	Correct	Incorrect	Total
Reference	y_j	x_j	m_j
Focal	y'_j	x'_j	m'_j
Total	n_j	n'_j	N_j

The Delta MH test statistic and variance have the following form:

$$DeltaMH = 2.35 \ln \frac{\left[\sum_{j=0}^{K} \frac{\left(y_j x'_j - y'_j x_j \right)}{N_j} \right]}{\sum_{j=0}^{K} \frac{y'_j x_j}{N_j}}$$

where y_j, x_j, y'_j, and x'_j are the frequency counts of cells of the 2 × 2 tables, and N_j is the total n for the cells.

The critical values of the ETS categorizations are 1.00 and 1.50 on the delta scale for categories A (negligible DIF), B (slight to moderate DIF), and C (moderate to severe DIF). Specifically, if the absolute value of delta is smaller than 1.00, the item is categorized as A. If the absolute value of delta is larger than or equal to 1.50, the item is classified as C. Otherwise items are categorized as B. In both the A and C categories, statistical significance is set at the 5% level for a single item.

The standardized mean difference (SMD) was computed for CR items. The SMD statistic (Dorans, Schmitt, and Bleistein, 1992) compares the mean scores of reference and focal groups, after adjusting for proficiency differences. The SMD was also evaluated for statistical significance and, in terms of practical significance, a moderate amount of DIF, for or against the focal group, is represented by an SMD with an absolute value between 0.10 and 0.19, inclusive; a large amount of DIF is represented by an SMD with an absolute value of 0.20 or greater.

Tables 5.3 and 5.4 present the results of the DIF analyses for the items included on the Fall 2018 and Spring 2019 operational forms, respectively. In these analyses, male students and white students were considered the reference groups for gender and ethnicity, respectively. The female students were the focal group for gender and the Black and Hispanic students were the focal groups for ethnicity. DIF analyses are performed when there is a minimum of 200 students in the focal group.

Table 5.3. Differential Item Functioning Analysis Results—Fall 2018

Content Area	Group	Core Form	*n*-Count	Dichotomous Items					Polytomous Items				
				A	B	B–	C	C–	A	B	B–	C	C–
English I	M/F	A	88/75	0	0	0	0	0	0	0	0	0	0
	W/B	A	91/33	0	0	0	0	0	0	0	0	0	0
	W/H	A	91/28	0	0	0	0	0	0	0	0	0	0
English II	M/F	A	1,326/1,205	39	1	0	0	0	0	0	0	0	0
	W/B	A	1,540/545	40	0	0	0	0	0	0	0	0	0
	W/H	A	1,540/277	37	1	2	0	0	0	0	0	0	0
Algebra I	M/F	A	2,770/2,528	37	1	2	0	0	3	0	1	0	0
	W/B	A	3,490/1,026	37	1	2	0	0	3	0	1	0	0
	W/H	A	3,490/448	37	2	1	0	0	4	0	0	0	0
Algebra II	M/F	A	273/272	37	3	2	1	1	2	0	0	0	0
	W/B	A	417/44	0	0	0	0	0	0	0	0	0	0
	W/H	A	417/50	0	0	0	0	0	0	0	0	0	0
Geometry	M/F	A	66/72	0	0	0	0	0	0	0	0	0	0
	W/B	A	111/8	0	0	0	0	0	0	0	0	0	0
	W/H	A	111/10	0	0	0	0	0	0	0	0	0	0
Biology	M/F	A	667/648	31	1	0	0	0	9	0	0	0	0
	W/B	A	851/231	32	0	0	0	0	8	0	1	0	0
	W/H	A	851/115	0	0	0	0	0	0	0	0	0	0
	M/F	B	604/583	37	0	1	1	1	5	0	0	0	0
	W/B	B	743/239	38	0	1	0	1	4	0	0	0	1
	W/H	B	743/101	0	0	0	0	0	0	0	0	0	0
Physical Science	M/F	A	17/22	0	0	0	0	0	0	0	0	0	0
	W/B	A	26/1	0	0	0	0	0	0	0	0	0	0
	W/H	A	26/12	0	0	0	0	0	0	0	0	0	0

Note. Classifications with a negative sign (“–”) favor the reference group, while classifications with no sign favor the focal group.
DIF contrast groups: M/F = male versus female; W/B = White versus Black; and W/H = White versus Hispanic.

Table 5.4. Differential Item Functioning Analysis Results—Spring 2019

Content Area	Group	Core Form	*n*-Count	Dichotomous Items					Polytomous Items				
				A	B	B–	C	C–	A	B	B–	C	C–
English I	M/F	C	2,926/2,865	28	0	0	0	0	6	0	0	0	0
	W/B	C	4,413/708	27	0	1	0	0	6	0	0	0	0
	W/H	C	4,413/411	26	0	2	0	0	6	0	0	0	0
	M/F	D	2,529/2,597	31	1	0	0	0	4	0	0	0	0
	W/B	D	3,908/712	26	1	4	0	1	3	0	0	0	1
	W/H	D	3,908/290	31	0	0	0	1	4	0	0	0	0
English II	M/F	C	16,407/16,249	33	0	1	0	0	3	0	0	0	0
	W/B	C	23,717/4,799	33	0	1	0	0	3	0	0	0	0
	W/H	C	23,717/2,087	34	0	0	0	0	3	0	0	0	0
	M/F	D	13,744/13,645	29	1	0	0	0	5	0	0	0	0
	W/B	D	20,322/3,692	28	0	2	0	0	5	0	0	0	0
	W/H	D	20,322/1,694	30	0	0	0	0	5	0	0	0	0
Algebra I	M/F	C	16,084/16,011	32	0	0	0	2	6	0	0	0	0
	W/B	C	23,136/4,733	33	0	0	0	1	5	0	1	0	0
	W/H	C	23,136/2,141	33	0	1	0	0	6	0	0	0	0
	M/F	D	13,307/13,197	32	0	0	0	2	7	0	0	0	0
	W/B	D	19,497/3,449	31	0	1	0	2	7	0	0	0	0
	W/H	D	19,497/1,695	33	0	1	0	0	7	0	0	0	0
Algebra II	M/F	C	3,801/4,281	32	0	0	0	0	6	0	1	0	0
	W/B	C	6,369/616	32	0	0	0	0	5	0	0	0	2
	W/H	C	6,369/426	32	0	0	0	0	5	0	2	0	0
	M/F	D	3,486/4,030	30	2	0	0	1	7	0	0	0	0
	W/B	D	6,055/511	30	0	1	0	2	6	0	1	0	0
	W/H	D	6,055/397	31	0	1	0	1	7	0	0	0	0
Geometry	M/F	C	887/952	29	2	3	0	0	7	0	0	0	0
	W/B	C	1,591/64	0	0	0	0	0	0	0	0	0	0
	W/H	C	1,591/87	0	0	0	0	0	0	0	0	0	0
	M/F	D	827/892	37	2	0	0	0	5	0	0	0	0
	W/B	D	1,498/58	0	0	0	0	0	0	0	0	0	0
	W/H	D	1,498/74	0	0	0	0	0	0	0	0	0	0
Biology	M/F	A	17,282/16,627	32	0	0	0	0	9	0	0	0	0
	W/B	A	24,552/5,115	32	0	0	0	0	8	0	1	0	0
	W/H	A	24,552/2,081	32	0	0	0	0	9	0	0	0	0
	M/F	B	13,230/13,330	36	2	1	0	1	4	1	0	0	0
	W/B	B	19,448/3,746	38	0	2	0	0	4	0	0	0	1
	W/H	B	19,448/1,698	39	0	0	0	1	4	0	0	0	1

Content Area	Group	Core Form	*n*-Count	Dichotomous Items					Polytomous Items				
				A	B	B–	C	C–	A	B	B–	C	C–
Physical Science	M/F	A	1,230/1,128	27	0	2	1	0	10	0	0	0	0
	W/B	A	2,032/120	0	0	0	0	0	0	0	0	0	0
	W/H	A	2,032/54	0	0	0	0	0	0	0	0	0	0

Note. Classifications with a negative sign ("–") favor the reference group, while classifications with no sign favor the focal group. DIF contrast groups: M/F = male versus female; W/B = White versus Black; and W/H = White versus Hispanic.

5.2.4. Summary

The item analyses provided in this chapter show that the MO EOC assessments have sound psychometric properties. For example, *p*-values show that MO EOC assessment items measure achievement across a broad range of difficulty. In addition, item discrimination values show that most items are appropriately correlated with the total test score and thus contribute to distinguishing between lower-performing and higher-performing students. Also, very few students omitted items during testing. The low percentage of students omitting items provides evidence that the test is a power test of the students' skills and not a speeded test. Additionally, DIF analyses conducted on gender and ethnicity help address construct-irrelevant variance, which presents a serious threat to the validity of inferences made from achievement test scores.

5.3. Overview of IRT Procedures

This section details the item response theory (IRT) item calibration procedures, the analysis procedures that relate to equating test scores from different forms, and information about the scaling of the MO EOC assessments. The purposes of scaling and equating are to maintain the consistency of the MO EOC assessments score scales over time and ensure that the achievement levels are applied consistently from year to year.

For the five English and Mathematics content areas (English I, English II, Algebra I, Algebra II, and Geometry), two core forms (C and D) were equated on the base-year scale using the data from both Fall 2018 and Spring 2019 administrations. The base-year scale of two core forms (A and B) was established using the data from both Fall 2017 and Spring 2018 administrations. The base-year scale was used in the standard setting workshop to get the scaled score cut points to be used to categorize students into achievement levels in 2018.

For the two Science content areas (Biology and Physical Science), two core forms (A and B) for Biology and one core form (A) for Physical Science were calibrated as the base-year scale using the data from Spring 2019 administrations. The core forms were used in the standard setting workshop to get the scaled score cut points to be used to categorize students' achievement levels. Subsequently, the reporting score scale was developed and used to report student scale scores on both administrations after the approval of the DESE.

For the two Social Studies content areas (Government and American History), the stand-alone field tests (SAFT) were administered in Spring 2019 and those items were used to construct the operational core forms for the Fall 2019 and Spring 2020 administrations. Because these were field tests, students' Social Studies scores were not available in Fall 2018 and Spring 2019.

The scaling and equating methods used for the English and Mathematics assessments are consistent with the methods used in the past. Detailed procedures for scaling and equating for the 2014–2015 administration are provided in the *2014–2015 MO EOC Technical Report*. Technical Report located on the DESE website at http://dese.mo.gov/college-career-readiness/assessment/assessment-technical-support-materials.

The following section begins with a description of the IRT models used for equating, followed by an overview of the scaling procedures, focusing on the five English and Mathematics operational tests. Note that calibration and scaling information for Science tests is included, but not equating because it was the first year of operational administration.

5.4. Calibration

Generally, item calibration is the process of estimating the parameters (such as item difficulty) for each item on an assessment so that all items are placed on a common scale. This section (a) introduces the Rasch model as the item response model used for MO EOC, (b) evaluates several IRT assumptions, and (c) summarizes item statistics for the operational MO EOC tests.

5.4.1. IRT Model Specification

The IRT models make several strong assumptions related to dimensionality, local independence, model-data fit, and item parameter invariance. The resulting inferences from any application of IRT depend on the degree to which the underlying assumptions are met.

Rasch IRT model scale is "a method for obtaining objective, fundamental, linear measures from stochastic observations of ordered category responses" (Linacre, 2006, p. 10). One feature of the Rasch model that distinguishes it from classical test theory is the placement of estimates of a person's ability and item difficulty on the same scale. The Rasch model expresses the probability of a correct response to an item as a function of the ability of the person and the difficulty of the item. In the Rasch model, the probability of a correct response to item *i*, given θ, is

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}},$$

where θ = latent trait, or ability, level, and b_i= the difficulty parameter for item *i*.

Masters (1982) developed the partial credit model (PCM) as an extension of the Rasch model to handle polytomous items, or items that allow for partially correct responses (e.g., open-ended items).

For an item with possible scores ranging from zero to *J*, the probability of obtaining score *j* on item *i*, given θ, is

$$P_{ij}(\theta) = \frac{e^{\sum_{k=0}^{j}(\theta - d_{ik})}}{\sum_{x=0}^{J} e^{\sum_{k=0}^{j}(\theta - d_{ik})}},$$

where d_{ij} is the difference between the overall item difficulty, b_i, and the step parameter γ_{ij} for level *j* of item *i*, and the sum of step parameters is zero across all levels of item *i*.
Winsteps software version 3.90.2 (Linacre, 2015) was used to perform the scaling and equating for the MO EOC assessments during the Spring 2019 administration. Winsteps is designed to produce a single scale by jointly analyzing data from students' responses to both dichotomous items and polytomous items. The dichotomous items were calibrated using the Rasch model (Rasch, 1960; Wright & Stone, 1979), while the partial credit model (Masters, 1982) was used to calibrate the polytomous items.

5.4.2. Checking IRT Assumptions

Because the Rasch and PCM models were the basis of all calibration, scoring, and scaling analyses associated with the MO EOC, the validity of the inferences from these results depends on the degree to which the assumptions of the model are met and how well the model fits the test data. Therefore, it is important to check these assumptions. This section evaluates the dimensionality of the data and item fit. It should be noted that only operational items were analyzed here since they are the basis of student scores.

5.4.2.1. Local Independence

Local independence refers to a response to an item that is not affected by other items. The IRT model assumes that a response to an item is only affected by the item's difficulty and a student's ability. Examples of violation of local independence are:

- Response to an item depends on the response to a prior item—such as, derive a value from item A then use item A's response to solve item B's equation. If both items A and B are answered incorrectly, it is not clear whether it is due to a lack of knowledge and skill on item A or item B.
- Other items on the test give away the answer to item A—this is referred to as clueing in test development.

The MO EOC handles local independence during item development and form construction process. All MO EOC tasks are written so that individually scored items do not have interdependency. During test construction, forms are reviewed for clueing and other properties to ensure the independence of test items.

Statistical analysis of local independence is assessed using Yen's (1984) Q3. Given the item residuals (d_i),

$$d_i = X_i - E\left(X_i|\hat{\theta}\right),$$

Q3 is calculated as the Pearson correlation of the item residuals

$$Q_{3,ij} = r_{d_i d_j},$$

where d_i and d_j are the item residuals for items i and j, respectively. Chen and Thissen (1997) state that a maximum absolute Q3 value greater than 0.20 is commonly used to identify the presence of local dependence.

Table 5.5 presents the number of unique item pairs, maximum absolute Q3 values, and number of flagged item pairs by content area. Across all subjects, less than 1% of the Q3 statistics were flagged for being greater than an absolute value of 0.20. Given the sample dependency of the Q3 statistic, the specific values, and qualitative review conducted during item development and form construction, it is assumed that the one item flagged in Algebra II, Biology, and Physical Science occurred by chance. For English I and English II, the items flagged were the dimension scores obtained from the total writing prompt score. Thus, an interdependence of dimension scores after accounting for the primary dimension is expected. Considering the qualitative review and quantitative analyses, the results suggest that the assumption of local independence is tenable.

Table 5.5. Number of Item Pairs, Maximum Absolute Q3 Statistic, and Number of Flagged Items by Content Area

Content Area	N	Maximum Absolute Value	# Flags
Algebra I	4,465	0.16	0
Algebra II	4,371	0.27	1
Geometry	4,753	0.15	0
English I	4,278	0.50	4
English II	3,741	0.48	4
Biology	2,556	0.24	1
Physical Science	780	0.23	1

5.4.2.2. Unidimensionality

Unidimensionality refers to a test that measures a single latent trait. The model fit statistics are indicators of dimensionality: Because the IRT models assume data are unidimensional, items have to be unidimensional to fit into the IRT model. Winsteps manual (Linacre, 2006) explains

> The Rasch model forces its estimates to be additive. Misfit means that the reported estimates, though effectively additive, provide a distorted picture of the data…The fit analysis is a report of how well the data accord with those additive measures. So a MnSq >1.5 suggests a deviation from unidimensionality in the data, not in the measures. So the unidimensional, additive measures present a distorted picture of the data. (p. 581–582)

5.4.2.3. Item Fit

Item fit refers to how well an item fits the calibration model (e.g., Rasch and PCM). It is usually a statistical chi-square, representing the difference between the observed score (i.e., actual student responses to items) and the expected score (i.e., what the model predicts students with a certain ability should be getting on the item). In order to check individual item fit for the MO EOC, Winsteps provided two informative statistics:

- **Infit:** An information-weighted fit statistic that is more sensitive to unexpected behavior affecting responses to items near the student's ability level

- **Outfit:** An outlier-sensitive fit statistic that is more sensitive to unexpected behavior by persons on items far from the student's ability level

Both infit and outfit provides mean-square fit (MNSQ) and its *z*-standardized statistics (ZSTD). Mean-square of infit and outfit has the expected value of the fit statistics as 1.0. Values that are less than 1.0 indicate that observations are too predictable (i.e., data over-fit the model), while values greater than 1.0 indicate unpredictability (i.e., data under-fit the model). The acceptable value of the infit and outfit is from 0.5 to 1.5 according to the Winsteps manual.
Table 5.5 shows the summary statistics of infit and outfit mean-squares, including the mean, standard deviation, and minimum and maximum values. The number of items within the range of [0.5, 1.5] is also reported.

Most items were within the range for infit mean square and all items had acceptable outfit mean-square. Some items had infit and outfit values greater than 1.5 or less than 0.5. These items were reviewed by assessment specialists, who confirmed that the items were correct in terms of the answer key and content.

Table 5.6. Summary of MNSQ INFIT and OUTFIT

Content Area	Mean	SD	Min	Max	*k*	*k* | [0.5, 1.5]
INFIT						
English I	1.00	0.10	0.82	1.50	93	92
English II	1.00	0.11	0.77	1.37	87	87
Algebra I	1.00	0.11	0.78	1.32	95	95
Algebra II	0.99	0.13	0.78	1.43	94	94
Geometry	0.98	0.10	0.78	1.25	98	98
Biology	0.99	0.11	0.76	1.26	72	72
Physical Science	1.00	0.10	0.84	1.20	40	40
OUTFIT						
English I	0.99	0.16	0.62	1.80	93	92
English II	1.00	0.18	0.52	1.44	87	87
Algebra I	1.00	0.18	0.59	1.40	95	95
Algebra II	0.99	0.20	0.70	1.91	94	92
Geometry	0.98	0.16	0.61	1.41	98	98
Biology	1.00	0.19	0.56	1.51	72	69
Physical Science	0.99	0.13	0.73	1.31	40	40

Note. *k* = number of items.

5.5. Equating

This section describes analysis procedures that relate to equating test scores from different MO EOC forms. In large-scale assessment programs, different item sets can be used on test forms either within years, across years, or both. Linking the scores from these different test forms puts the form scores on a common scale of measurement and ensures that all forms for a given content-area test provide comparable scores. This means that students will not have an unfair advantage or disadvantage simply because the test form they took was easier or more difficult than the test forms taken by other students.

Most of MO EOC tests (Physical Science and American History being the exception) currently use two sets of operational core forms (called A and B, for example) per administration. Those core forms are coupled with multiple field test form variants. The scaled scores of English and Mathematics content areas on both core A and B were established after the Spring 2018 administration using the Rasch and PCM IRT model. The base scale was created using the data from the Fall 2017 and Spring 2018 administrations. Equating analyses on scaled scores on both core C and D have been conducted for the Spring 2019 test administration to link test scores across administrations so that student performance is always expressed on the common scaled score metric. Since Science test scale scores were established at the Spring 2019 baseline test administration and standard setting in the summer of 2019, they are not included in the equating analysis procedures. Social studies tests are excluded from equating analysis since they were administered as a stand-alone field test for Spring 2019.

5.5.1. Post-equating

Two core forms were administered during the Spring 2019 window for MO EOC assessments in English, Mathematics, and Biology. When multiple core forms are administered in the same window using a spiraling technique, it can be assumed that the student groups taking each form are randomly equivalent and a common metric is created during calibration using a concurrent calibration. Once the common metric is developed, an external anchor nonequivalent groups design is used to equate across administrations. In this equating design, no assumption is made about the equivalence of the examinee groups taking different test forms (i.e., naturally occurring groups are assumed). Comparability is instead evaluated by using a set of anchor items (also called equating items), assuming they perform in the same way in both groups and can, thus, accurately measure the differences in the two groups.

The post-equating procedure used eligible external anchor sets designated by content experts during test construction. The external anchor sets were intended to be "mini-test" meaning that they matched the content and statistical characteristics of the core forms. After anchor stability was assessed through the multiple methods described in the next section, the remaining anchor items were used to derive item parameters of the operational items in the core forms using a fixed common item parameter calibration method. Finally, given the calibrated item parameters, the RSS tables were then derived.

There are multiple administration formats of the MO EOC assessments. Except for students who need format accommodations, students are administered the MO EOC assessments via online testing platform, Nextera®. Small numbers of students take the assessment using an alternate testing format (Braille, Large Print, or Paper/Pencil).

5.5.2. Evaluation of Anchor Items

The evaluation of the anchor item set and determination of the final anchor set are the most critical steps of the equating procedure. The anchor item set was originally selected in the form construction stage by assessment development experts. The following four statistical analyses are proposed to assess the stability and validity of the anchor items.

5.5.2.1. Position Shift Analysis

Questar compared the anchor item positions in the current test book to the one in the previous administration. To the extent possible, anchor items were sequenced as close as possible to the position in which they were field tested. Anchor items that were sequenced more than five positions away from where they were tested were reviewed.

5.5.2.2. Item-total Correlation

If the correlation coefficient was less than 0.2, the anchor item was flagged and removed. The item-total correlation is provided in the point-measure correlation for all observations (PTMA) column of the Winsteps calibration output file; it utilizes the ability estimates as the criterion score, not the raw scores.

5.5.2.3. Robust Z Analysis

Robust Z analyses were performed by comparing each equating item's new item parameter estimates from the calibration step with the banked item parameter estimates from the previous administration. Robust Z is calculated as the differences between the banked and new difficulties using the following formula:

$$Robust\ Z_i = \frac{Rasch_i^{diff} - Median(Rasch_i^{diff})}{0.74\ \cdot IQR}$$

where i indexes anchor items, and IQR stands for the inter quartile range. As recommended by Huynh and Meyer (2010), if any of the below criteria are not met, any anchor item with an absolute Robust Z value of 1.645 was flagged and removed. However, if all of the criteria are met, no anchor items were removed, regardless of Robust Z value.

- The ratio of the standard deviations of the previous and current year Rasch difficulties is in the 0.9 and 1.1 range.
- The correlation between the previous year and current year Rasch difficulties is greater than 0.95.

5.5.2.4. D-squared Analysis

In order to determine if each item from the anchor set performs similarly to when it was previously administered, the individual item characteristic curves (ICCs) from the previous and current administrations are compared.

The linked item parameter estimates are used to calculate a weighted squared deviation of the current ICC from the previous ICC, across the range of ability (i.e., θ) and under a hypothetical normal distribution for θ. For a given item i, that quantity, called d-squared, is given by

$$d_{\mathrm{i}}^2 = \sum_{\mathrm{k}} \left\{ \left[\mathrm{Pr}_{\mathrm{i,new}}(\theta_{\mathrm{k}}) - \mathrm{Pr}_{\mathrm{i,old}}(\theta_{\mathrm{k}}) \right]^2 \cdot \mathrm{g}(\theta_{\mathrm{k}}) \right\},$$

where i indexes anchor items, k indexes quadrature points for θ, $\mathrm{Pr}_{\mathrm{i,new}}(\theta_{\mathrm{k}})$ is the probability of a correct response to item i under the current calibration, while $\mathrm{Pr}_{\mathrm{i,old}}(\theta_{\mathrm{k}})$ is the same quantity under the previous calibration, and $\mathrm{g}(\theta_{\mathrm{k}})$ are weights for the quadrature points. A fixed criterion on this metric ($d_{\mathrm{i}}^2 \geq 0.05$) is used for flagging items to be considered for removal from linking.

The following stages were be taken:

1 Before the iterative process starts, the initial linking is performed using all of the eligible anchor items. The initial linking coefficients are obtained through the Stocking-Lord (SL) method.

2 These processes are repeated until the iteration is less than or equal to five:

- 2.1 For each anchor item, d_i^2 will be calculated.
- 2.2 If the largest d_i^2 value of any anchor items is greater than 0.05, it is removed from the anchor set.
  - 2.2.1 Linking is performed with the newly reduced anchor set.
  - 2.2.2 The processes from stage 2.1 are repeated.
- 2.3 If d_i^2 values of all anchor items are less than 0.05, the iteration is stopped, and the current anchor set is considered the final one.
  - 2.3.1 Go to Stage 3.

3 The list of anchor items with the flag identifications is produced as the result of the above stages.

Note that the second stage is repeated five times at most or until the largest $d_i^2 < 0.05$ is reached, whichever is smaller. Once the lists of flags from those four analyses are completed, any flagged anchor items will be considered as outliers and excluded from the linking process.

The anchor evaluation results in Table 5.6 contain information for the linking items excluded from the equating analyses. The table includes the following for each linking item: position change between prior and current item sequence position, item-ability correlation (PTMA), Robust Z statistic, and D squared statistics. The reason for exclusion from the equating is added in the last column.

Table 5.7. Anchor Evaluation Results

Content	Item ID	Position Shift	PTMA	Robust Z	D Squared	Reason Excluded
English I	MOE116365	**18**	0.25	-1.76	0.002	Position Shift
English II	MOE116294	**15**	0.40	-3.25	0.007	Position Shift
	MOE116295	**19**	0.41	-4.22	0.002	Position Shift
	MOE116300	**14**	0.37	-2.03	0.003	Position Shift
	MOE116355	**14**	0.51	-2.99	0.006	Position Shift
	MOE116358	**14**	0.42	-1.91	0.002	Position Shift
	MOE216218	**-8**	0.27	-2.33	0.004	Position Shift
	MOE216224	-8	0.14	-0.35	0.000	PTMA
	MOE216776	**6**	0.51	-2.25	0.002	Position Shift
Algebra II	MOA21619	5	0.46	-6.84	0.013	Robust Z
	MOA216218	0	0.49	2.15	0.003	Robust Z
	MOA216427	-2	0.38	-3.84	0.004	Robust Z
	MOA216432	**-6**	0.59	0.01	0.000	Position Shift
	MOA216441	-1	0.48	1.85	0.003	Robust Z
	MOA216445	-2	0.27	-5.64	0.007	Robust Z
	MOA216497	3	0.47	1.81	0.002	Robust Z
	MOA21677	**-6**	0.36	-3.15	0.002	Position Shift
	MOA2169	0	0.51	-2.42	0.002	Robust Z
Geometry	MOG16362	0	0.12	-1.07	0.001	PTMA
	MOG16789	**21**	0.55	-2.54	0.007	Position Shift

5.6. Scaling

Scaling transforms IRT ability estimates (i.e., theta) to a reporting scale that is easier for test score users to understand and enables legitimate and meaningful comparisons. Scaled scores should facilitate proper score interpretations while minimizing misinterpretations and unwarranted inferences. This is often done by attaching content meaning to the scores (Peterson, Kolen, and Hoover, 1989). Many state assessments add content meaning by anchoring one or more of a test's performance-level cut scores to known scaled-score values.

The *Standards for Educational and Psychological Measurement* (APA, AERA, NCME, 2014, p. 102) provide the following guidelines regarding scaling:

- Standard 5.1. Test users should be provided with clear explanations of the characteristics, meaning, and intended interpretation of scaled scores, as well as their limitations.
- Standard 5.2. The procedures for constructing scales used for reporting scores and the rationale for these procedures should be described clearly.
- Standard 5.3. If there is sound reason to believe that specific misinterpretations of a score scale are likely, test users should be explicitly cautioned.

5.6.1. Scale Process

Scaling was conducted to transform theta into scale scores. Raw score is the total number of points a student earns on an assessment for answering items correctly. Theta has a one-to-one relationship with raw score. However, the raw score alone does not present an adequate picture of test performance because it can be interpreted only in terms of a particular set of items. In order to allow for meaningful comparisons of scores across test forms and administrations, theta, which has been equated, are transformed to the assessment's common scale through the scaling process.

Individual student scores are reported as scaled scores. However, the student scores are initially estimated as IRT ability estimates. Scaled scores are preferable to theta estimates for reporting purposes because theta values have negative and decimal values. By transforming theta values to scaled scores, all reported values are positive integers with no decimals and are thought to be easier for students and parents to understand. When new test forms are administered in future administrations, scale scores can be used to compensate for any differences in the difficulty of the items and allow direct comparisons of student performance.

Based on the results of the policy review meeting and subsequent approval by the State Board of Education, the scale scores for the MO EOC were obtained following the required properties below:

a. The lowest obtainable scale score (LOSS) value exists and is 325
b. The highest obtainable scale score (HOSS) will vary across years
c. A scale score of 400 is designated as the scale score cut of the Proficient performance level
d. One standard deviation of student ability in the baseline year administration equals 15 scale score points
e. The corresponding average and standard deviation on the ability distribution are used to calculate the slope and intercept of the transformation from the ability estimate to the scale score

The baseline scales for each of five English or Mathematics tests were built on the two core operational forms using the data from both Fall 2017 and Spring 2018 administrations. From the Spring 2019 administration, the two new operational core forms are scaled on the baseline scale. Additionally, the baseline scales for Science tests are built for two Biology core forms and one Physical Science form from the Spring 2019 administration.

5.6.2. Slopes and Intercepts

To produce the theta-to-scale score (and RSS) conversion tables, a linear transformation was applied to theta estimates. The following formula was used to obtain the slope and intercept for the transformation function:

$$sc(y) = \left[\frac{sc(y_2) - sc(y_1)}{\theta_2 - \theta_1}\right] y + \left\{sc(y_1) - \left[\frac{sc(y_2) - sc(y_1)}{\theta_2 - \theta_1}\right]\theta_1\right\}$$

where θ_2 and θ_1 are theta estimates that correspond to the cut score points, and $sc(y_2)$ and $sc(y_1)$ are scale score points. This formula was adopted from Kolen and Brennan (2004, p.337).

Table 5.7 presents the slopes and intercepts for the RSS linear transformation along with the theta and scale score cut points.

Table 5.8. Theta to Scale Score Transformation with Slopes and Intercepts

Content Area	Basic		Proficient		Advanced		Slope	Intercept
	Theta	SS	Theta	SS	Theta	SS		
English I	-1.0722	384	0.0000	400	1.0389	415	14.8947	400.0000
English II	-1.0611	384	0.0111	400	1.3778	420	14.5594	399.8384
Algebra I	-1.0667	389	-0.2389	400	0.5056	409	12.7533	403.0468
Algebra II	-0.9444	388	0.0833	400	1.0000	411	11.6912	399.0261
Geometry	-1.0667	387	-0.0722	400	0.9556	414	13.3786	400.9659
Biology	-0.4971	381	0.9073	400	1.7293	411	13.5020	387.7496
Physical Science	-0.4558	382	0.6102	400	1.5789	417	17.0553	389.5928

5.6.3. RSS Conversions Tables

Appendix E provides the RSS conversions tables for Fall 2018 and Spring 2019 for each core forms in a content area. Once the slope and intercept were estimated in the base form year, they were used to create the RSS conversion tables by using the linear equation presented above for the base form and future administrations.

Chapter 6: Performance Level Setting

One purpose of assessment is to establish clear guidelines for educational decision-making. By assigning meaning to test scores, standard setting allows policymakers, administrators, teachers, parents, and students to make statements about the level of proficiency of individual students and groups of students.

- In 2018, standard setting was conducted for English I, English II, Algebra I, Algebra II, and Geometry.
- In 2019, standard setting was conducted for Biology and Physical Science.

The new Missouri Learning Standards were approved in April 2016 and implemented in a sequence of over three years. The 2019 standard setting established cut points for Biology and Physical Science. The report with more detailed information on the 2019 standard setting meeting is presented in Appendix C. Also, consult the *2018 Standard Setting Report* for the English Language Arts and Mathematics as an additional reference. The cut points for Government and American History will be established in 2020.

6.1. 2019 Standard Setting

In 2016, the Missouri State Board of Education approved new Missouri Learning Standards for Science; these standards were implemented in 2017–2018. The MAP began assessing these standards in 2018–2019. As part of a multi-phased standard setting, DESE sought to establish new cut points for MAP Science assessments which: (a) reflect the new Missouri Learning Standards, (b) link students' scores on the MAP to the state's expectations for students in each performance level, and (c) are well articulated across grades and courses.

The performance levels for MAP are designed to indicate students' knowledge of the skills listed in the Missouri Learning Standards. The performance levels are *Below Basic, Basic, Proficient*, and *Advanced*.

6.1.1. Standard Setting Methodology

A total of 47 Missouri educators engaged in the Bookmark Standard Setting Procedure (Lewis, Mitzel, & Green, 1996; Lewis, Mitzel, Mercado, & Schulz, 2012) to recommend cut scores. This method has been used on assessments in Missouri and across the nation, including for English Language Arts and Mathematics of the MAP. There were 13 participants for the Grade 5 assessment, 11 for Grade 8[10], 12 for Biology, and 11 for Physical Science.

Participants studied the updated Missouri performance level descriptors (PLDs) and Missouri Learning Standards to review the knowledge, skills, and abilities expected of students in each performance level. Each performance level was associated with a level of mastery of the Missouri Learning Standards. Participants then discussed the content-based expectations for students at the threshold of each performance level (e.g., a student who is just *Proficient*).

[10] The grades 5 and 8 assessment Standard Setting was not conducted by Questar.

Participants studied *ordered item booklets* (OIBs) that comprised collections of operational test items that were ordered by difficulty. A separate OIB was created for each test, and items' difficulty values were based on students' performance on the test items. Participants studied the OIBs to understand the knowledge and skills measured by the tests.

On May 2nd, a benchmark panel of Missouri educators and stakeholders was convened to help DESE select appropriate benchmarks for the four science tests. This panel also provided contextual information to accompany the benchmarks when they were ultimately used at the performance level setting workshop.

Unlike the Grades 5 and 8 assessments, a parallel NAEP assessment program does not exist for the EOC assessments. NAEP has a Grade 12 assessment, but it does not provide statewide results and not all states participate in the administration of the assessment. At the benchmark panel meeting held on May 2nd, the panel endorsed the use of NAEP for the Grade 5 and Grade 8 assessments, but it could not find suitable external benchmarks for the EOC assessments. The benchmark panel developed recommendations for benchmarks for the two EOC assessments, but it did not express a significant amount of confidence in the recommendations. Just as with Grades 5 and 8, a band of ±1 CSEM was used to create a benchmark band, loosely associated with a *Proficient* range, and this range was reflected in participants' OIBs. However, participants were told that the information was provided simply for the participants' information, and that the participants' recommendation would not necessarily be consistent with the benchmark band. Participants engaged in three rounds of individual judgments and group discussions. In each round, participants recommended cut scores by considering the content-based expectations for students in each performance level, and then identifying the sets of items in their OIBs which best represented these expectations. By placing bookmarks, participants recommended cut scores on the test scale.

Between rounds, participants were shown feedback (e.g., median bookmarks, impact data). The committees' median judgments were taken as their recommendations. For Rounds 1 and 2, feedback was provided to each individual panel and only shown the results for their specific program. After Round 3 ratings were completed, all participants were shown the results across all grade levels and programs to allow the entire panel to discuss the results and evaluate the consistency of their specific program with the other Science assessments.

After the Round 3 discussion, table leaders convened to examine the recommendations. As needed, table leaders recommended adjustments to promote articulation among the performance standards across grades. When examining the cut scores, the table leaders recommended adjustments to two cut scores to improve the articulation of the entire system of performance standards. The table leaders noted that the percent of students classified as *Below Basic* in Grade 8 was markedly lower than observed in other grades, so it recommended adjusting its recommended bookmark to 17 from 15. Similarly, the table leaders reported that the percentage of Biology students classified as *Advanced* was unexpectedly high, especially when compared to the other grades, and recommended adjusting their bookmark to 84 from 83.

Several table leaders noted that the percentage of students classified as *Proficient* and above in Grade 8 was markedly higher than that observed in the other three grades. These table leaders noted that this pattern was unexpected; however, the table leaders could not make a consensus recommendation to adjust the cut score. After the standard setting, this cut score was adjusted by one conditional standard error of measurement (CSEM) value: the CSEM quantifies the amount of statistical noise around a point on the test scale. Accordingly, the Grade 8 *Proficient* cut score was adjusted to 510 from 501, a change of +1 CSEM. This adjustment was made to promote articulation across grades while still honoring the judgments of Missouri educators at the standard setting.

Table 6.1 shows the recommended cut scores for the MAP Science tests, plus the associated impact data using Spring 2019 administration data. Impact data are the percentages of students who would be classified in each performance level if the cut scores were applied to students' scores.

Table 6.1. Recommended Cut Scores and Associated Impact Data for the MAP Science Tests

Content Area	Grade	Recommended Cut Scores			% Students by Level Based on Spring 2019				
		Basic	Proficient	Advanced	Below Basic	Basic	Proficient	Advanced	Prof. + Adv.
Science	5	275	310	344	26.0%	31.1%	29.3%	13.5%	42.8%
	8	468	510	537	20.4%	35.7%	25.9%	18.1%	43.9%
	PS	382	400	417	19.2%	42.9%	28.2%	9.7%	37.9%
	BIO	381	400	411	16.8%	43.8%	24.5%	15.0%	39.5%

6.1.2. Statistical Standard Error Estimates

For each recommended cut point, the conditional standard error of measurement (CSEM) was calculated at that point on the operational test form. Using these CSEM values, the corresponding cut score plus or minus 1 and 2 CSEM intervals was calculated. The resulting cut scores and the corresponding impact data are provided in Table 6.2.

Table 6.2. Cut Scores Adjusted by CSEM Intervals and Associated Impact Data for the MAP Science Tests

Interval	Performance Level	Cut Scores Adjusted by CSEM				Associated Impact Data (%)			
		5	8	PS	BIO	5	8	PS	BIO
+2 CSEM	Below Basic	--	--	--	--	44.45	36.43	40.80	32.58
	Basic	297	490	392	389	33.25	37.82	41.26	48.99
	Proficient	332	528	410	409	18.80	18.99	16.41	14.27
	Advanced	368	555	429	421	3.50	6.76	1.53	4.17
	Prof + Adv	--	--	--	--	22.30	25.75	17.94	18.44
+1 CSEM	Below Basic	--	--	--	--	34.53	27.66	28.84	24.96
	Basic	286	479	387	385	33.27	37.79	42.88	46.14
	Proficient	321	519	405	404	24.86	22.98	24.64	19.97
	Advanced	356	546	423	416	7.34	11.57	3.65	8.93
	Prof + Adv	--	--	--	--	32.20	34.55	28.29	28.90
As Recommended	Below Basic	--	--	--	--	26.05	20.38	19.17	16.77
	Basic	275	468	382	381	31.11	35.71	42.92	43.77
	Proficient	310	510	400	400	29.32	25.86	28.20	24.49
	Advanced	344	537	417	411	13.53	18.05	9.71	14.97
	Prof + Adv	--	--	--	--	42.85	43.91	37.91	39.46
-1 CSEM	Below Basic	--	--	--	--	18.82	14.54	10.01	10.79
	Basic	264	457	377	377	27.53	32.37	37.74	39.73
	Proficient	299	501	395	396	31.35	27.34	34.31	25.99
	Advanced	332	528	410	406	22.30	25.75	17.94	23.50
	Prof + Adv	--	--	--	--	53.65	53.09	52.25	49.49
-2 CSEM	Below Basic	--	--	--	--	12.91	9.99	3.44	5.75
	Basic	253	446	371	373	23.44	28.21	33.29	32.53
	Proficient	288	492	390	391	30.45	27.25	31.89	25.72
	Advanced	320	519	404	402	33.20	34.55	31.38	35.99
	Prof + Adv	--	--	--	--	63.65	61.80	63.28	61.71

Chapter 7: Studies of Reliability and Construct-related Validity

7.1. Introduction

Evidence of internal structure of test scores is the center of validity arguments. This chapter provides studies of reliability and construct-related validity evidence focusing on the consistency of the internal assessment structure.

7.2. Reliability

DESE is required to ensure that the instruments used to measure student achievement for school accountability provide reliable results. As Standard 2.0 of the *Standards for Educational and Psychological Testing* states "Appropriate evidence of reliability/precision should be provided for the interpretation for each intended score use" (p. 42). This chapter provides evidence that scores from the MO EOC assessments measure student achievement in a reliable manner[11] and that the size of the measurement error associated with reported test scores is reasonable[12], especially at the Proficient cut score.

7.2.1. Defining Reliability

According to the *Standards for Educational and Psychological Testing* and consistent with the measurement literature, reliability is defined two different ways:

> First, the term has been used to refer to the reliability coefficients of classical test theory, defined as the correlation between scores on two equivalent forms of the test, presuming that taking one form has no effect on performance on the second form. Second, the term has been used in a more general sense, to refer to the consistency of scores across replications of a testing procedure, regardless of how this consistency is estimated or reported. (p. 33)

In general, reliability refers to the consistency of student test scores, or the extent to which an assessment yields the same results repeatedly. Reliability considers random error, which results from outside influences that can affect a student's score. An assessment that produces highly consistent, stable results (i.e., mostly free from random error) is considered highly reliable. The less random error, the more reliable the test scores are. The more reliable the assessment scores are, the more consistent a student's test scores will be if the student takes a replicated version of the test (i.e., a test that has different items but that covers the same topics using the same number of items per topic). Reliability can be estimated via the correlation of scores on forms assumed to be parallel (equivalence reliability), from test-retest data (stability reliability), or from a single test administration (internal consistency reliability).

[11] **Standard 2.3:** For each total score, subscore, or combination of scores that is to be interpreted, estimates of relevant indices of reliability/precision should be reported (AERA, APA, NCME, 2014, p. 43).
[12] **Standard 2.13:** The standard error of measurement, both overall and conditional (if reported), should be provided in units of each reported score (AERA, APA, NCME, 2014, p. 45).

7.2.2. Reliability Coefficient

Classical test theory (CTT) provides a means for quantifying reliability. In CTT, an observed measurement, such as test score (*X*) is defined as a composite of true score (*T*) and an associated random error component (*E*):

$$X = T + E.$$

The definitions and assumptions in CTT lead to several important properties. For example, it can be demonstrated that observed score variance equals the sum of (1) the variance in true scores—true individual differences in the attribute being measured, and (2) the variance from random fluctuations due to the imperfections in the measurement process (error variance).

$$\sigma_x^2 = \sigma_t^2 + \sigma_e^2$$

Normally, a covariance term is required when adding variances, but it is not in this case as true scores and errors are assumed to be uncorrelated in CTT. The reliability coefficient expresses the consistency of test scores as the ratio of true-score variance to total observed-score variance.

$$\rho_{x_1 x_2} = \frac{\sigma_t^2}{\sigma_x^2}$$

Reliability coefficients theoretically range from 0.0 to 1.0, although the extremes are never achieved in applied testing programs. Larger coefficients are more desirable because they indicate that test scores are less influenced by random error. If all test score variance was true, the scores would be perfectly consistent and the index would equal 1.0. The index would be 0.0 if none of the test score variance was true. Such scores would be pure random noise (i.e., all measurement error).

7.2.3. Estimating Reliability

The reliability of a specific test cannot be directly estimated from the equation above. Although several different reliability indices exist, an industry-standard index for describing internal consistency reliability based on a single test administration is coefficient alpha (Cronbach, 1951), which provides an estimate of reliability that is mathematically equivalent to the average of all possible split-half reliability estimates computed by the Rulon method. For a test consisting of *p* items, in which the item scores Y_j are summed to get a total score *X*, coefficient alpha is computed as follows:

$$\alpha = \left(\frac{p}{1-p}\right)\left(1 - \frac{\sum_{j=1}^{p} \sigma_{Y_j}^2}{\sigma_X^2}\right)$$

7.2.4. Interpretation Considerations

The coefficient alpha indicates the internal consistency of the responses over a set of items measuring an underlying trait, in this case, academic achievement in the MO EOC content tests. As an internal consistency index, it can be conceptualized as indicating the extent to which an exchangeable set of items from the same domain would result in a similar rank ordering of students.

Relative error is reflected by coefficient alpha. Further, coefficient alpha is only sensitive to random errors due to the sampling of items. It does not take into account other random sources of error (e.g., variations associated with the linking process; daily fluctuation in student health and behavior, the testing environment; rater inconsistency).

7.3. Reliability Evidence

Reliability evidence for the 2018–2019 MO EOC assessments includes the following:

- Internal consistency
- Standard error of measurement (SEM) for raw scores
- Conditional standard error of measurement (CSEM) for Scale Scores
- Classification accuracy and consistency
- Rater agreement (presented in Section 4.4)

7.3.1. Standard Error of Measurement (SEM) for Raw Scores

No test provides a perfect measure of a student's ability because all tests have a known standard error of measurement (SEM). The SEM represents the amount of variability that can be expected in a student's test score because of the inherent imprecision of the test. For example, if the student were tested again with a new test of comparable difficulty, he or she would likely obtain a slightly different score. The expected range for this new score is provided as a standard error (SE) and gives an indication of the margin of error for the reported scale score.

7.3.1.1. Traditional SEMs and Traditional Confidence Intervals (CIs)

The SEM is defined as the standard deviation of the distribution of observed scores for students with identical true scores. The standard deviation is a measure of the dispersion of the observed scores; for the normal distribution, about 32 percent of observations are more than one standard deviation above or below the mean.

The SEM formula:

$$SEM = \sigma_X \sqrt{1 - \alpha}$$

indicates that the value of the SEM depends on both the reliability coefficient and the standard deviation of test scores.

SEMs allow statements regarding the overall precision of test scores. SEMs help place "reasonable limits" (Gulliksen, 1950) around observed scores through construction of an approximate score band or confidence intervals (CIs). These bands are constructed by taking the observed scores, X, and adding and subtracting a multiplicative factor of the SEM. As an example, students with a given true score will have observed scores that fall between +/-1 SEM about two-thirds of the time.

7.3.1.2. <u>Reliabilities and SEMs by Student Subgroup</u>

Separate analyses were performed for each EOC content area. The tables in Appendix J provides the reliabilities and SEMs for the total population and for select student subgroups. For each table, the effect size, reliability, and SEM are reported for each group provided there were at least 50 students in the group.

An effect size is reported within each group, provided minimal sample size requirements are met. The effect size is a measure of how much the scores of two groups of students differ from each other. It is based on score standard deviations, and calculated using Cohen's *d* equation:

$$d = \frac{\bar{X}_F - \bar{X}_R}{\hat{\sigma}_X},$$

where the numerator is the difference in average scores between a focal and a reference group, and the denominator is an estimate of total score standard deviation. In this case, the standard deviations across groups were pooled to generate the standard deviation estimate.

An effect size of 1.0 is equivalent to a difference of one standard deviation. An effect size of 0.8 is considered *large*; an effect size of 0.5 is considered *medium*; an effect size of 0.2 is considered *small*. Effect sizes are also reported whenever the reference and focal groups each have a minimum of 50 students.

Following MO EOC program convention, the reference groups are gender = Male, ethnicity = White, LEP status = no, IEP status = no, Migrant status = no, FRL status = no, Title 1 status = no, and Accommodations status = no.

Tables 7.1 and 7.2 provide the reliability and SEM along with the number of student and average and standard deviation of raw scores by core forms.

Table 7.1 Overall Reliability and SEM by Core Form—Fall 2018

Content Area	Core Form	*n*-Count	Mean Raw Score	SD Raw Score	Reliability	SEM
English I	A	162	25.82	8.69	0.86	3.22
English II	A	2,509	23.18	9.90	0.89	3.35
Algebra I	A	5,271	21.26	10.70	0.92	3.08
Algebra II	A	545	23.04	10.60	0.92	3.04
Geometry	A	137	31.80	9.22	0.91	2.83
Biology	A	1,315	26.20	11.67	0.92	3.28
	B	1,186	26.20	11.67	0.93	3.07
Physical Science	A	39	--	--	0.76	--

Note. Due to the small sample size, Physical Science is excluded from calculation.

Table 7.2 Overall Reliability and SEM by Core Form—Spring 2019

Content Area	Core Form	*n*-Count	Mean Raw Score	SD Raw Score	Reliability	SEM
English I	C	5,767	28.68	8.02	0.85	3.09
	D	5,098	28.68	8.02	0.83	3.27
English II	C	32,562	28.53	8.32	0.87	3.03
	D	27,311	28.53	8.32	0.84	3.28
Algebra I	C	32,031	22.31	9.68	0.89	3.17
	D	26,451	22.31	9.68	0.88	3.35
Algebra II	C	8,062	24.93	9.30	0.88	3.26
	D	7,501	24.93	9.30	0.88	3.20
Geometry	C	1,837	22.03	8.97	0.87	3.20
	D	1,718	22.03	8.97	0.86	3.32
Biology	A	33,873	30.30	10.16	0.90	3.15
	B	26,533	30.30	10.16	0.91	3.09
Physical Science	A	2,286	27.17	8.79	0.86	3.25

7.3.1.3. Interpretation Considerations

The SEM approach only provides a single numerical estimate for constructing confidence intervals for examinees regardless of their score level. In reality, such confidence intervals vary according to a student's score. Consequently, care should be taken using the SEM for students with extreme scores. Because test reliabilities and standard deviations are group specific, the same is true for SEMs and CIs. For the MO EOC, the SEM approach is calculated using raw scores, and as such, the resulting confidence interval bands are in the raw-score metric.

7.3.2. Conditional Standard Error of Measurement (CSEM) for Scale Scores

7.3.2.1. CSEMs and Conditional CIs

According to the *Standards for Educational and Psychological Testing*, Standard 2.14 states:

> When possible and appropriate, conditional standard errors of measurement should be reported at several score levels unless there is evidence that the standard error is constant across score levels. Where cut scores are specified for selection or classification, the standard errors of measurement should be reported in the vicinity of each cut score. (p. 46)

This section describes the calculation of the CSEMs. As noted below, the CSEMs for each scale score are presented in Appendix E and the CSEMs at the Proficient cut are presented in Table 7.3.

Rasch-based CSEMs are also used for the MO EOC assessments. CSEMs also allow statements regarding the precision of individual test scores by helping derive reasonable limits around observed scaled scores through construction of approximate score bands, referred to as conditional confidence intervals (CIs). Any given test will have CSEMs that vary as a function of the scaled scores. This makes the CSEM especially useful in characterizing measurement precision around a score level used for decision making, such as a cut score used for identifying students who meet a given performance standard.

MO EOC CSEMs come from the Winsteps program and are based on the principle of statistical information. The CSEM at any given point on the ability (θ, theta) continuum is defined as the reciprocal of the square root of the test information function derived from the Rasch scaling model. In the formula, $CSEM(\hat{\theta})$ is the conditional standard error of measurement, and $I(\hat{\theta})$ is the test information function:

$$CSEM\ (\hat{\theta}) = \frac{1}{\sqrt{I(\hat{\theta})}}.$$

Test information depends on the sum of the corresponding information functions for the test items. Item information depends on each item's unique conditional item score variance as determined from its difficulty parameters and conditional item score variance. The formula provides the CSEMs on the Rasch ability (θ) metric.

7.3.2.2. CSEMs at the Proficient Cut

CSEMs are useful for characterizing measurement precision in the neighborhood of score levels used for decision making, such as cut scores at various achievement levels. The CSEMs for the Proficient cut scores for the MO EOC assessments are presented in Table 7.3. The CSEM values were 5 for English I and English II. The CSEM values for the Mathematics content areas were 3 through 5 in Fall and Spring. The CSEM for Biology and Physical Science were 4 or 5 for the Fall and Spring administrations. The CSEM values for Government and American History were not included in the table due to the lack of student reporting this year. CSEMs for the other scale scores are reported in Appendix E. Note that CSEMs are smaller in the middle of the score distribution than at the extremes. This pattern is expected for CSEMs based on item response theory (IRT).

Table 7.3. CSEM at Proficient Cut Score

Test Period	Content Area	SS Cut*	CSEM
Fall 2018	English I	400	5
	English II	400	5
	Algebra I	400	4
	Algebra II	400	3
	Geometry	400	5
	Biology	400	4
	Physical Science	400	5
Spring 2019	English I	400	5
	English II	400	5
	Algebra I	400	4
	Algebra II	400	3
	Geometry	400	5
	Biology	400	4
	Physical Science	400	5

Note. Appendix E contains the CSEM at each scale score.

7.3.3. Classification Accuracy and Consistency

The accuracy and consistency of classifying students into achievement levels are critical components of a standards-based reporting framework (Livingston & Lewis, 1995). For the MO EOC tests, students are classified into one of four achievement levels. Questar conducted classification accuracy and consistency analyses to determine the statistical accuracy and consistency of the classifications. This section explains the methodologies used to assess the reliability of classification decisions and gives the results of these analyses.

7.3.3.1. Classification Accuracy and Consistency as a Measure of Reliability

Classification accuracy refers to the accuracy of decisions (e.g., the accuracy of students' assignments to achievement levels), or the extent to which decisions would agree with those that would be made if each student could somehow be tested with all possible versions of the assessment, which implies that the scores did not contain any measurement error. Accuracy must be estimated, because errorless test scores do not exist.

Consistency measures the extent to which classifications based on test scores match the classifications based on scores from a second, parallel form of the assessment that is equal in difficulty and covers the same content as the form the students actually took. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are administered to the same group of students. In operational testing programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and consistency of classifications based on a single administration of a test.

The Livingston and Lewis (1995) technique addresses the single administration of a test by making use of "true scores" in the classical test theory sense. A true score is the score that would be obtained if a test had no measurement error. True scores cannot be observed and so must be estimated. The estimated true scores are used to categorize students into their "true" classifications.

As described in the Livingston and Lewis (1995), using the BB-CLASS for PC software (Brennan, 2004), a four-by-four contingency table of accuracy was calculated for each grade, where cell $[i, j]$ represented the estimated proportion of students whose true score fell into classification i (where i = 1 to 4) and observed score fell into classification j (where j = 1 to 4). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

To calculate consistency, true scores were used to estimate the joint distribution of classifications on two independent, parallel test forms. Following the same statistical procedures, a new four-by-four contingency table was calculated for each grade and populated by the proportion of students who would be categorized into each combination of classifications according to the two (hypothetical) parallel test forms. Cell $[i, j]$ of this table represented the estimated proportion of students whose observed score on the first form would fall into classification i (where i = 1 to 4) and whose observed score on the second form would fall into classification j (where j = 1 to 4). The sum of the diagonal entries (i.e., the proportion of students categorized by the two forms into exactly the same classification) signified overall consistency.

In addition to the overall consistency, Cohen's (1960) coefficient K (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance, is calculated using the following formula:

$$K = \frac{(Observed\ agreement) - (Chance\ agreement)}{1 - (Chance\ agreement)} = \frac{\Sigma_i C_{ii} - \Sigma_i C_{i.} C_{.i}}{1 - \Sigma_i C_{i.} C_{.i}}$$

where

$C_{i.}$ is the proportion of students whose observed achievement level would be level i (where i = 1–4) on the first hypothetical parallel form of the test;
$C_{.i}$ is the proportion of students whose observed achievement level would be level i (where i = 1–4) on the second hypothetical parallel form of the test; and
C_{ii} is the proportion of students whose observed achievement level would be level i (where i = 1–4) on both hypothetical parallel forms of the test.

Because K is corrected for chance, its values are lower than other consistency estimates. Based on the four-by-four contingency tables used to estimate the overall accuracy and consistency, the classification accuracy and consistency conditional on achievement level are also evaluated.

Consistency conditional on achievement level is conceived as the ratio between the proportion of correct classifications at the selected achievement level and the proportion of all the students classified into that level.

Accuracy conditional on achievement level is conceived in a similar manner, except that in the consistency table where both row and column marginal sums are the same, the accuracy table uses the sum based on estimated status as the total for computing accuracy conditional on achievement level.

For some testing situations where the greatest concern may be decisions around achievement level thresholds, the primary concern is distinguishing between students who are proficient and those who are not yet proficient. In this case, accuracy at the Basic/Proficient threshold is critically important, which summarizes the percentage of students who are correctly classified either above or below the particular cutpoint. To evaluate decisions at specific cut scores, the same four-by-four contingency tables are used.

The accuracy index at the cut score is computed as the sum of the proportions of correct classifications around this selected cut score.

The consistency at a specific cut score is obtained in a similar way but involves computing the sum of the proportions of consistent classifications around this selected cut score.

7.3.3.2. Decision Accuracy and Consistency Results

Results of the DAC analyses described above for both Fall 2018 and Spring 2019 are provided in Table 7.4 and 7.5. These tables include overall accuracy indices with consistency indices displayed in parentheses next to the accuracy values, as well as overall kappa values. Overall ranges for accuracy (0.73–0.81), consistency (0.63–0.75), and kappa (0.49–0.62) indicate that the vast majority of students were classified accurately and consistently with respect to measurement error and chance.

Accuracy and consistency values conditional on achievement level are also given in the next columns. For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, the conditional accuracy value is 0.82 for Below Basic for Fall 2018 English I. This figure indicates that among the students whose true scores placed them in this classification, 82% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.70 indicates that 70% of students with observed scores in the Below Basic would be expected to score in this classification again if a second, parallel test form was taken.

Table 7.4. Summary of Decision Accuracy (and Consistency) Results Fall 2018—Overall and Conditional on Achievement Level

			Conditional on Achievement Level			
Content Area	Overall	Kappa	Below Basic	Basic	Proficient	Advanced
English I	0.73 (0.63)	0.49	0.82 (0.70)	0.68 (0.59)	0.70 (0.61)	0.81 (0.68)
English II	0.77 (0.68)	0.55	0.85 (0.77)	0.70 (0.61)	0.78 (0.70)	0.79 (0.61)
Algebra I	0.76 (0.68)	0.57	0.88 (0.82)	0.68 (0.57)	0.61 (0.50)	0.86 (0.78)
Algebra II	0.77 (0.69)	0.58	0.88 (0.81)	0.71 (0.61)	0.67 (0.56)	0.86 (0.77)
Geometry	0.80 (0.73)	0.59	0.78 (0.58)	0.75 (0.64)	0.71 (0.62)	0.90 (0.85)
Biology	0.81 (0.73)	0.62	0.88 (0.82)	0.80 (0.74)	0.66 (0.54)	0.85 (0.74)
Physical Science	--	--	--	--	--	--

Note. Due to the small sample size, Physical Science is excluded from calculation.

Table 7.5. Summary of Decision Accuracy (and Consistency) Results Spring 2019—Overall and Conditional on Achievement Level

			Conditional on Achievement Level			
Content Area	Overall	Kappa	Below Basic	Basic	Proficient	Advanced
English I	0.74 (0.64)	0.50	0.80 (0.65)	0.71 (0.61)	0.71 (0.63)	0.82 (0.71)
English II	0.78 (0.69)	0.53	0.81 (0.67)	0.73 (0.64)	0.80 (0.74)	0.80 (0.66)
Algebra I	0.73 (0.64)	0.52	0.84 (0.75)	0.68 (0.58)	0.61 (0.50)	0.85 (0.76)
Algebra II	0.75 (0.65)	0.52	0.81 (0.68)	0.73 (0.64)	0.71 (0.61)	0.83 (0.72)
Geometry	0.76 (0.66)	0.52	0.82 (0.69)	0.74 (0.65)	0.73 (0.64)	0.82 (0.69)
Biology	0.78 (0.69)	0.56	0.84 (0.73)	0.80 (0.74)	0.66 (0.55)	0.83 (0.73)
Physical Science	0.76 (0.67)	0.51	0.82 (0.69)	0.76 (0.69)	0.73 (0.63)	0.79 (0.63)

Tables 7.6 through 7.7 provide accuracy and consistency estimates for the Fall 2018 and Spring 2019 MO EOC tests at each cutpoint, respectively, as well as false positive and false negative decision rates. A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. Whereas, a false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut. The accuracy and consistency indices at the Basic/Proficient threshold range from 0.88–0.93 and

0.83–0.89. The false positive and false negative decision rates at the Basic/Proficient threshold range from 3–6% and 3–6%. These results indicate that nearly all students were correctly classified with respect to being above or below the Basic/Proficient cutpoints.

Table 7.6. Summary of Decision Accuracy (and Consistency) Results Fall 2018—Conditional on Cut Score Point

	Test	English I	English II	Algebra I	Algebra II	Geometry	Biology	Physical Science
Below Basic/Basic	Accuracy (Consistency)	0.93 (0.90)	0.91 (0.87)	0.92 (0.88)	0.92 (0.89)	0.98 (0.97)	0.92 (0.89)	--
	False Positive	0.03	0.04	0.04	0.04	0.01	0.03	--
	False Negative	0.05	0.05	0.04	0.04	0.02	0.04	--
Basic/ Proficient	Accuracy (Consistency)	0.88 (0.83)	0.90 (0.85)	0.91 (0.87)	0.91 (0.88)	0.92 (0.89)	0.93 (0.89)	--
	False Positive	0.06	0.06	0.05	0.05	0.03	0.04	--
	False Negative	0.06	0.05	0.04	0.04	0.04	0.03	--
Proficient/ Advanced	Accuracy (Consistency)	0.92 (0.89)	0.97 (0.95)	0.93 (0.91)	0.94 (0.92)	0.91 (0.87)	0.96 (0.94)	--
	False Positive	0.05	0.02	0.04	0.03	0.05	0.03	--
	False Negative	0.03	0.01	0.03	0.02	0.05	0.02	--

Note. Due to the small sample size, Physical Science is excluded from calculation.

Table 7.7. Summary of Decision Accuracy (and Consistency) Results Spring 2019—Conditional on Cut Score Point

	Test	English I	English II	Algebra I	Algebra II	Geometry	Biology	Physical Science
Below Basic/Basic	Accuracy (Consistency)	0.95 (0.92)	0.95 (0.92)	0.92 (0.89)	0.94 (0.91)	0.93 (0.90)	0.94 (0.91)	0.92 (0.89)
	False Positive	0.02	0.02	0.03	0.02	0.03	0.03	0.03
	False Negative	0.03	0.03	0.05	0.04	0.04	0.04	0.05
Basic/ Proficient	Accuracy (Consistency)	0.88 (0.84)	0.89 (0.84)	0.89 (0.85)	0.89 (0.84)	0.89 (0.84)	0.90 (0.86)	0.89 (0.84)
	False Positive	0.06	0.05	0.05	0.06	0.06	0.05	0.06
	False Negative	0.06	0.06	0.05	0.06	0.06	0.05	0.05
Proficient/ Advanced	Accuracy (Consistency)	0.91 (0.88)	0.94 (0.92)	0.92 (0.88)	0.92 (0.89)	0.94 (0.91)	0.94 (0.91)	0.95 (0.93)
	False Positive	0.05	0.04	0.05	0.05	0.04	0.04	0.03
	False Negative	0.03	0.02	0.03	0.03	0.02	0.02	0.02

7.4. Construct-related Validity Evidence

This section summarizes the validity evidence as it relates to the purpose and intended use of the MO EOC test results (refer to Section 1.2). Validity evidence based on the internal structure of the MO EOC assessments is then provided through a correlational analysis of MO EOC assessment content clusters. References to specific standards are provided where appropriate.

7.4.1. Internal Structure

The item analyses shown in Appendix D revealed that the MO EOC assessments have sound psychometric properties. The p-value ranges were sufficiently broad, indicating that the items measure achievement across a broad range of difficulty. Item-test correlations, indicators of item discrimination, are also provided. Almost all items had acceptable discrimination values (i.e., discrimination values > 0.15). Some extremely difficult items had low discrimination values that were likely attenuated by their difficulty.

Empirical investigation of DIF strengthens the validity evidence related to score interpretations for students in particular groups by evaluating potential sources of construct-irrelevant variance. DIF results might be better considered as internal—structure validity evidence. Statistical analyses results are provided in Chapter 4. The results indicated that none of the PE/WP items were flagged for DIF and that either no or very few SR items were flagged for DIF across subjects and administrations. Standard 1.13[13] pertains to the relationships between the parts of the test. Because the MO EOC assessments measure student performance in several content areas, it is important to study the pattern of relationships among the content domains and clusters.

Tables 7.8–7.14 summarize correlation coefficients among test domains and clusters for English I, English II, Algebra I, Algebra II, Geometry, Biology, and Physical Science. Because the correlation coefficients will be affected by the limited number of items measuring each domain, the correlation coefficient between two content standard clusters may be artificially low because of measurement error. Therefore, the correlations are corrected for attenuation. The formula for the correlation coefficient statistically corrected for attenuation (r_{ca}) is Spearman's formula

$$r_{ca} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}$$

where r_{xy} is the correlation between content clusters, r_{xx} is the reliability of one content cluster, and r_{yy} is the reliability of the other content cluster.

The tables report the Pearson correlations below the diagonal, the correlations corrected for attenuation above the diagonal, and Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells. The corrected correlations between clusters within each assessment are strong (> 0.80), with many correlations exceeding 0.90. The disattenuated correlations greater than 1.00 that indicate that measurement error is not randomly distributed are reported as 1.00.

[13] **Standard 1.13:** If the rationale for a test score interpretation for a given use depends on premises about the relationships among test items or among parts of the test, evidence concerning the internal structure of the test should be provided (AERA, APA, NCME, 2014, p. 26–27).

Each content area test is comprised of two or more content clusters measuring a single construct or dimension. These results suggest that the cluster scores are appropriately related to each other. Therefore, the results provide evidence that a unidimensional construct is measured on each of the MO EOC assessments supporting the validity of the test construct.

Table 7.8. Correlation Coefficients Between Domains and Clusters—English I

Admin	Cluster	Core Form	Reading Literary Texts	Reading Informational Texts	Writing
Fall 2018	Reading Literary Texts	A	0.69	0.98	1.00
	Reading Informational Texts	A	0.69	0.72	1.00
	Writing	A	0.67	0.70	0.60
Spring 2019	Reading Literary Texts	C	0.70	1.00	0.93
	Reading Informational Texts	C	0.64	0.57	0.95
	Writing	C	0.66	0.61	0.73
	Reading Literary Texts	D	0.64	0.98	0.93
	Reading Informational Texts	D	0.63	0.65	0.92
	Writing	D	0.59	0.59	0.63

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

Table 7.9. Correlation Coefficients Between Domains and Clusters—English II

Admin	Cluster	Core Form	Reading Literary Texts	Reading Informational Texts	Writing
Fall 2018	Reading Literary Texts	A	0.68	0.99	0.97
	Reading Informational Texts	A	0.68	0.71	0.95
	Writing	A	0.71	0.71	0.80
Spring 2019	Reading Literary Texts	C	0.66	0.99	0.95
	Reading Informational Texts	C	0.66	0.67	0.96
	Writing	C	0.67	0.68	0.76
	Reading Literary Texts	D	0.67	0.94	0.89
	Reading Informational Texts	D	0.59	0.59	0.93
	Writing	D	0.61	0.61	0.72

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

Table 7.10. Correlation Coefficients Between Domains and Clusters—Algebra I

Admin	Cluster	Core Form	Algebra	Functions	Number/Quantity and Statistics
Fall 2018	Algebra	A	0.85	0.97	0.92
	Functions	A	0.79	0.80	0.99
	Number/Quantity and Statistics	A	0.70	0.73	0.68
Spring 2019	Algebra	C	0.76	1.00	1.00
	Functions	C	0.80	0.80	1.00
	Number/Quantity and Statistics	C	0.66	0.68	0.56
Spring 2019	Algebra	D	0.77	1.00	1.00
	Functions	D	0.76	0.74	1.00
	Number/Quantity and Statistics	D	0.65	0.65	0.55

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

Table 7.11. Correlation Coefficients Between Domains and Clusters—Algebra II

Admin	Cluster	Core Form	Algebra	Functions	Number/Quantity and Statistics
Fall 2018	Algebra	A	0.88	1.00	1.00
	Functions	A	0.81	0.74	1.00
	Number/Quantity and Statistics	A	0.69	0.66	0.52
Spring 2019	Algebra	C	0.82	0.95	1.00
	Functions	C	0.71	0.68	1.00
	Number/Quantity and Statistics	C	0.65	0.58	0.44
	Algebra	D	0.84	0.98	1.00
	Functions	D	0.70	0.61	1.00
	Number/Quantity and Statistics	D	0.65	0.56	0.48

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

Table 7.12. Correlation Coefficients Between Domains and Clusters—Geometry

Admin	Cluster	Core Form	Congruence/ Similarity Coordinate Geometry & Circles	Geometry	Statistics and Probability
Fall 2018	Congruence/Similarity Coordinate Geometry & Circles	A	0.88	0.93	1.00
	Geometric Measurement Modeling	A	0.72	0.68	0.86
	Statistics and Probability	A	0.66	0.47	0.43
Spring 2019	Congruence/Similarity Coordinate Geometry & Circles	C	0.86	1.00	0.81
	Geometric Measurement Modeling	C	0.64	0.47	0.88
	Statistics and Probability	C	0.45	0.36	0.35
	Congruence/Similarity Coordinate Geometry & Circles	D	0.83	0.90	1.00
	Geometric Measurement Modeling	D	0.60	0.54	0.91
	Statistics and Probability	D	0.55	0.39	0.33

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

Table 7.13. Correlation Coefficients Between Domains and Clusters—Biology

Admin	Cluster	Core Form	From Molecules to Organisms: Structure and Process	Ecosystems: Interactions, Energy, and Dynamics	Heredity: Inheritance and Variation of Traits	Biological Evolution: Unity and Diversity	Earth and Human Activity
Fall 2018	From Molecules to Organisms: Structure and Process	A	0.68	0.98	1.00	0.94	1.00
	Ecosystems: Interactions, Energy, and Dynamics	A	0.71	0.77	0.99	0.98	1.00
	Heredity: Inheritance and Variation of Traits	A	0.73	0.73	0.70	0.95	1.00
	Biological Evolution: Unity and Diversity	A	0.69	0.76	0.71	0.79	1.00
	Earth and Human Activity	A	0.54	0.63	0.58	0.64	0.42
Spring 2019	From Molecules to Organisms: Structure and Process	A	0.62	0.96	1.00	0.94	0.98
	Ecosystems: Interactions, Energy, and Dynamics	A	0.64	0.73	0.95	0.96	1.00
	Heredity: Inheritance and Variation of Traits	A	0.66	0.67	0.69	0.92	0.96
	Biological Evolution: Unity and Diversity	A	0.64	0.71	0.66	0.75	1.03
	Earth and Human Activity	A	0.50	0.57	0.52	0.58	0.42
	From Molecules to Organisms: Structure and Process	B	0.67	1.00	1.00	1.00	1.00
	Ecosystems: Interactions, Energy, and Dynamics	B	0.78	0.80	0.99	0.99	1.00
	Heredity: Inheritance and Variation of Traits	B	0.75	0.78	0.79	0.96	1.00
	Biological Evolution: Unity and Diversity	B	0.73	0.77	0.74	0.75	1.00
	Earth and Human Activity	B	0.60	0.64	0.61	0.62	0.45

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

Table 7.14. Correlation Coefficients Between Domains and Clusters—Physical Science

Admin	Cluster	Core Form	Matter and Its Interactions	Motion and Stability: Forces and Interactions	Energy	Earth and the Universe
Fall 2018	Matter and Its Interactions	A	0.44	0.51	0.95	0.67
	Motion and Stability: Forces and Interactions	A	0.27	0.64	1.10	0.92
	Energy	A	0.38	0.54	0.37	1.33
	Earth and the Universe	A	0.28	0.46	0.51	0.39
Spring 2019	Matter and Its Interactions	A	0.66	0.90	0.95	1.02
	Motion and Stability: Forces and Interactions	A	0.62	0.72	0.97	1.09
	Energy	A	0.64	0.67	0.67	1.13
	Earth and the Universe	A	0.40	0.44	0.44	0.23

Note. Cronbach's coefficient alpha of the cluster scores on the diagonal in gray shaded cells.

7.4.2. Convergent and Divergent Validity

Convergent validity examines the extent to which theoretically related constructs are empirically related, whereas divergent validity examines the extent to which theoretically unrelated constructs are empirically unrelated. The Standards state the following regarding convergent and divergent validity: "Relationships between test scores and other measures intended to assess the same or similar constructs provide convergent evidence, whereas relationships between test scores and measures purportedly of different constructs provide discriminant evidence" (AERA, APA, & NCME, 2014, p. 16–17). The MO EOC assessments were designed to measure different constructs, as shown by both the standards they assess and the content coverage detailed in the test blueprints.

7.4.2.1. Pearson Correlations Among Assessments

Table 7.15 shows evidence of convergent and divergent validity. The data sets used for the analysis were drawn from the Spring 2019 operational test administration. The students in the data sets were merged using Missouri's unique student identification number. Any student who took at least two operational tests was included in the Pearson correlation coefficients between scale scores for Spring 2019.

Evidence of divergent validity is supported by the lower correlations between content areas that measure dissimilar constructs. For example, the correlations between English II and Algebra II (0.50) and between English I and Geometry (both 0.49) and between English II and Geometry (0.57) are in a range typical of achievement constructs that are positively related primarily by virtue of their relation to general school achievement.

Table 7.15 also provides more evidence of convergent validity. Evidence of convergent validity emerges when comparing correlations between the similar contents of Algebra I and Geometry (0.82), Algebra I and Algebra II (0.79), and Algebra II and Geometry (0.78). The correlation between Biology and Physical Science is 0.83. The scientific contents and skills on both tests could account for the higher correlation. Table 7.16 contains the n-counts of students who took multiple tests.

Table 7.15. Pearson Correlation among Assessments

Assessment	Algebra I	Algebra II	Geometry	English I	English II	Biology	Physical Science
Algebra I	1	0.79	0.82	0.60	0.70	0.73	0.68
Algebra II	0.79	1	0.78	0.54	0.50	0.61	0.47
Geometry	0.82	0.78	1	0.49	0.57	0.63	0.69
English I	0.60	0.54	0.49	1	0.69	0.74	0.71
English II	0.70	0.50	0.57	0.69	1	0.77	0.77
Biology	0.73	0.61	0.63	0.74	0.77	1	0.83
Physical Science	0.68	0.47	0.69	0.71	0.77	0.83	1

Table 7.16. N-Counts of Multiple Test Takers

Assessment	Algebra I	Algebra II	Geometry	English I	English II	Biology	Physical Science
Algebra I	--	199	106	5,745	12,761	16,577	1,221
Algebra II	199	--	179	451	10,234	5,673	62
Geometry	106	179	--	739	1,874	1,486	403
English I	5,745	451	739	--	51	565	2,040
English II	12,761	10,234	1,874	51	--	31,624	143
Biology	16,577	5,673	1,486	565	31,624	--	64
Physical Science	1,221	62	403	2,040	143	64	--

7.4.3. Additional Validity Evidence

Validity evidence related to other standards is described below.

- Standard 1.8[14] relates to the characteristics of the sample of test takers from which validity evidence is inferred. The sample of examinees from which the validity evidence for the MO EOC assessments is referred to this chapter. Appendix F summarizes the descriptive statistics of scale scores, and Appendix G summarizes the percentage distributions of students' achievement levels by demographic group.

- Standard 1.9[15] relates to human judgment at various points in the test development and reporting process. For the MO EOC assessments, human judgment was especially prevalent during the standard setting and cutpoint validation processes. When cut scores are critical to the interpretation of test results, the procedural validity of the processes used to establish those scores also should be addressed. Chapter 3 contains summary information about the standard setting procedures used for the MO EOC assessments. Overall, the panelists' feedback from both workshops indicated that they understood the process and were comfortable with their cut score recommendations. Human judgment is also a component of handscoring. From Spring 2008 through Spring 2010 and again in 2014–2015 and 2015–2016, PE/WPs were handscored. Chapter 6 contains detailed information about the processes involved with Questar's handscoring of the 2015–2016 PE/WPs, including scorer selection and training.

[14] **Standard 1.8:** The composition of any sample of test takers from which validity evidence is obtained should be described in as much detail as is practical and permissible, including major relevant sociodemographic and developmental characteristics (AERA, APA, NCME, 2014, p. 25).

[15] **Standard 1.9:** When a validation rests in part on the opinions or decisions of expert judges, observers, or raters, procedures for selecting such experts and for eliciting judgments or ratings should be fully described. The qualifications and experience of the judges should be presented. The description of procedures should include any training and instructions provided, should indicate whether participants reached their decisions independently, and should report the level of agreement reached. If participants interacted with one another or exchanged information, the procedures through which they may have influenced one another should be set forth (AERA, APA, NCME, 2014, p. 25).

- Standard 1.10[16] relates to the conditions under which the data used to support validity claims were collected. Chapter 5 contains information about how data were gathered in both the online and accommodated administrations, including the testing environment, materials distribution and security, Test Examiner training, student preparation, and allowable accommodations.

7.5. Summary

The validation process involves the ongoing collection of a variety of evidence to support the proposed test-score interpretations and uses. It is not an all-or-nothing property of a test; rather, evidence must be documented for a specific purpose and in the context of how the test scores will be interpreted and used. Much of the information contained in this technical report is validity evidence for the MO EOC assessments' stated purposes. This chapter provided a summary of the evidence presented elsewhere in the technical report and provided some additional types of validity evidence relevant to the content and internal structure of the assessments.

Post-administration test analyses supported the technical quality of the MO EOC assessments. Validity of score inferences is bolstered when test scores are consistent. Here, the reliabilities of the total test scores are very good, ranging from 0.83 to 0.93 across the content areas and administrations for the 2018–2019 test forms. The CSEMs were between five and nine scale score points at the Proficient cut scores. Additionally, DIF analyses conducted on gender and ethnicity help address construct-irrelevant variance, which presents a serious threat to the validity of inferences made from achievement test scores.

[16] **Standard 1.10:** When validity evidence includes statistical analyses of test results, either alone or together with data on other variables, the conditions under which the data were collected should be described in enough detail that users can judge the relevance of the statistical findings to local conditions. Attention should be drawn to any features of validation data collection that are likely to differ from typical operational testing conditions and that could plausibly influence test performance (AERA, APA, NCME, 2014, p. 26).

Chapter 8: Reporting and Results

8.1. Introduction

This chapter provides the information about MO EOC test score reporting and the descriptive summary of test score results for each of the seven MO EOC assessments from the Fall 2018 and Spring 2019 administrations, including the total raw scores, scale scores, and performance levels.

8.2. Types of Reports

The purpose of reporting assessment data is to communicate test results to students, parents, teachers, administrators, and other stakeholders. The MO EOC assessment reports provide useful information for determining the performance of students in a particular district, school, or classroom. These reports describe students' knowledge and skills regarding a set of expectations, allowing educators to determine specific instructional needs, measure student mastery toward postsecondary readiness, provide evidence of accountability for Missouri and national programs, and evaluate educational programs. Districts may also use locally designed assessments aligned to the Missouri Learning Standards to provide more detailed information for each student in specific content areas.

Questar delivers a General Research File (GRF) to DESE at the end of each test administration that contains all of the raw data collected for each administration. Questar also provides a *Guide to Interpreting Results* for DESE to post on their website that provides explanations of the ALD and Missouri Learning Standards strands for each content area, as well as samples of the Individual Student Report (ISR) and the Student Score Label with descriptions of the information they contain. ISRs are provided in the online assessment platform for all assessment windows. In addition, several csv files are also provided in the online assessment plat form. These include a roster and a percent of points earned. Student Score Labels are provided in hard copy to districts following each administration.

The MO EOC assessment score indicates that an individual student performs at the Below Basic, Basic, Proficient, or Advanced level in a given content area. ALDs provide details about the content expectations that students at each level meet or exceed. The scores are scaled in several ways: raw scores, scale scores (derived from the Rasch model), and achievement level (based on scale score cuts) that describe what students can do in terms of the content and skills assessed. These scores provide a way to compare test results with standards of academic performance. Subscale scores are not reported for the MO EOC assessments.

Missouri promotes the use of achievement level results, reporting them annually on each assessment at the student, school, district, and state levels. Individual student and average scale scores are also used, but they play a secondary role and are generally interpreted with reference to their distance from achievement level cut points.

To determine the achievement level scores, Questar converted each student's raw score points earned into a scale score (described in later in this chapter). The scale score determined the student's achievement level. Each achievement level represented standards of performance for each content area. Test results are reported for students as a whole and by student group, including gender, ethnicity, migrant status, free and reduced lunch (FRL) status, English language proficiency, Title I, IEP status, and accommodations used during testing. Scores are reported to schools and districts in annually published reports.

No stakes for teachers are attached to student-level scores by the state. Teachers are encouraged to consider student performance on the MO EOC assessments in determining course grades. DESE recommends that MO EOC scores account for at least 10 percent but not more than 25 percent of a student's grade in a course with a corresponding MO EOC assessment. Districts receive student scores on the MO EOC assessments within five business days after test administration, and DESE provides districts with "curved percentages" to assist teachers in appropriately considering EOC scores in determining course grades (http://dese.mo.gov/sites/default/files/asmt-eoc-curved-percentages.pdf). Teachers are counseled to interpret individual student scores only in the context of other assessment results and their own experience.

8.2.1. Individual Student Report (ISR)

The 2018–2019 Individual Student Report (ISR) provides information about performance on the MO EOC assessment, describing the results in terms of four levels of achievement in a content area. It is used for measuring an individual student's mastery toward postsecondary readiness for the content area. It is also used in instructional planning as a point of reference during parent-teacher conferences and for permanent record keeping. Teachers are informed that other sources of information should be used along with this report when determining the student's areas of strength or need.

On the report, achievement-level scores describe what students can do in terms of the CLEs for the content and skills assessed by the MO EOC assessment. A student at the Proficient or Advanced level has met the standard.

A sample of the ISR appears in Figure 8.1. A brief description of selected parts of the report is as follows:

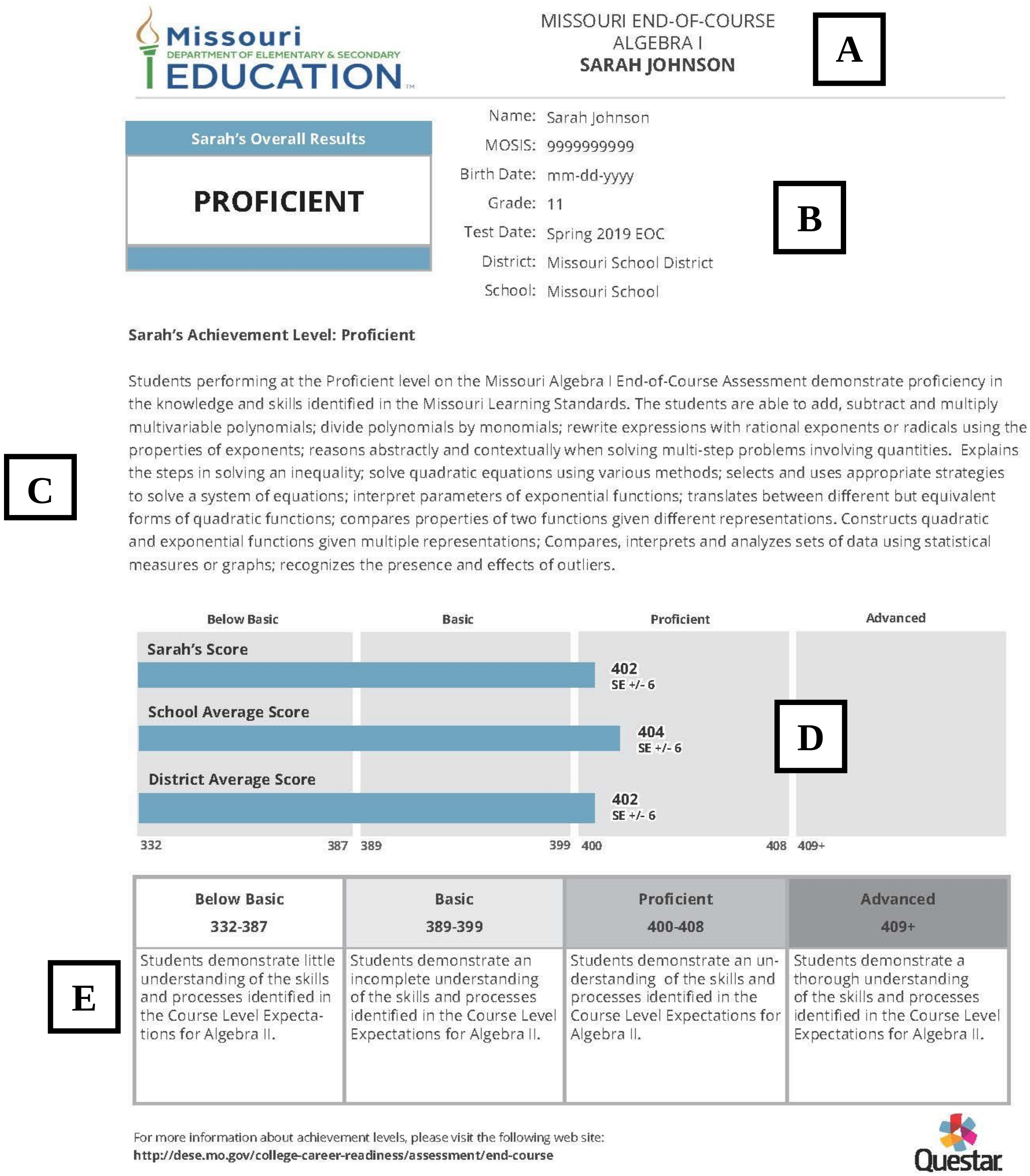
A. The heading of the ISR includes the content area for the results being presented. A separate report is produced for each content area tested.

B. The student information section contains the biographic data for the individual student taking the assessment. Identifying information including the MOSIS ID, date of birth, grade, test date, district, and school is listed, followed by the test period.

C. The narrative describes the student performance characteristics corresponding to the obtained achievement level. The text is specific to the content area tested. At the bottom of the narrative is a URL for a website that provides additional information for all of the achievement levels for the content area.

D. The mean scale scores for the student's school and district are displayed in the two rows below the student's individual results. The mean scale score, with an associated SE, and the bar graph provide a way to view the individual's results in contrast to the group's results for the content area during the same test period.

E. The individual student's results are presented numerically as a three-digit scale score with the SE. An accompanying bar graph to the right of the scale score illustrates the achievement level obtained by the student. Achievement levels (whether Below Basic, Basic, Proficient, or Advanced) are based on the scale score ranges listed beneath the Achievement Level heading in the table.



Missouri DEPARTMENT OF ELEMENTARY & SECONDARY EDUCATION

MISSOURI END-OF-COURSE
ALGEBRA I
SARAH JOHNSON

A

Sarah's Overall Results

PROFICIENT

Name: Sarah Johnson
MOSIS: 9999999999
Birth Date: mm-dd-yyyy
Grade: 11
Test Date: Spring 2019 EOC
District: Missouri School District
School: Missouri School

B

Sarah's Achievement Level: Proficient

C

Students performing at the Proficient level on the Missouri Algebra I End-of-Course Assessment demonstrate proficiency in the knowledge and skills identified in the Missouri Learning Standards. The students are able to add, subtract and multiply multivariable polynomials; divide polynomials by monomials; rewrite expressions with rational exponents or radicals using the properties of exponents; reasons abstractly and contextually when solving multi-step problems involving quantities. Explains the steps in solving an inequality; solve quadratic equations using various methods; selects and uses appropriate strategies to solve a system of equations; interpret parameters of exponential functions; translates between different but equivalent forms of quadratic functions; compares properties of two functions given different representations. Constructs quadratic and exponential functions given multiple representations; Compares, interprets and analyzes sets of data using statistical measures or graphs; recognizes the presence and effects of outliers.

E

Below Basic 332-387	Basic 389-399	Proficient 400-408	Advanced 409+
Students demonstrate little understanding of the skills and processes identified in the Course Level Expectations for Algebra II.	Students demonstrate an incomplete understanding of the skills and processes identified in the Course Level Expectations for Algebra II.	Students demonstrate an understanding of the skills and processes identified in the Course Level Expectations for Algebra II.	Students demonstrate a thorough understanding of the skills and processes identified in the Course Level Expectations for Algebra II.

For more information about achievement levels, please visit the following web site:
http://dese.mo.gov/college-career-readiness/assessment/end-course

Questar

Figure 8.1. Individual Student Report (ISR)

8.2.2. Student Score Label

The 2018–2019 Student Score Label provides a summary of a student's results on the MO EOC assessment. A separate label is produced for each content area tested. The individual label provides the student's biographic data, scale score, and achievement level. The labels have adhesive backing so they can be easily transferred onto the student record folders.

A sample label is shown in Figure 8.2. A brief description of selected parts of the label is as follows:

A. The top of the label shows the content area tested.
B. The student's name and identifying information are provided on the left side of the label as well as the student's scale score and achievement level. If a student has results for more than one content area, the next label is printed below the first one.

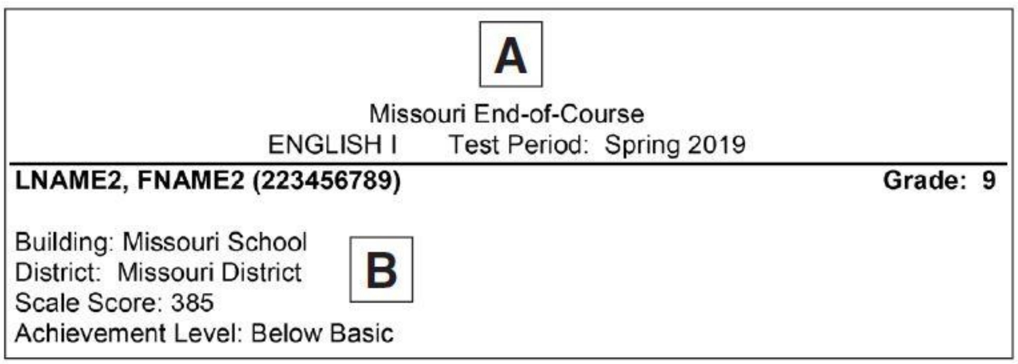


Figure 8.2. Student Score Label

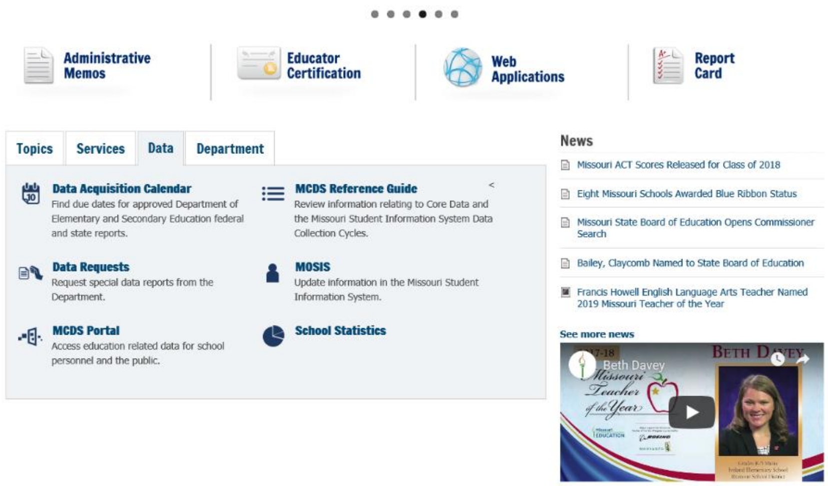
8.2.3. Missouri Comprehensive Data System (MCDS) Portal

8.2.3.1. Purpose and Use

For the first two years of the MO EOC assessment administration, summary-level EOC results were available to school district personnel in a set of standard reporting configurations through DESE's Crystal Reporting system. Reporting options included administrative reports, adequate yearly progress (AYP) reports, achievement level reports, content standard reports, and item analysis reports.

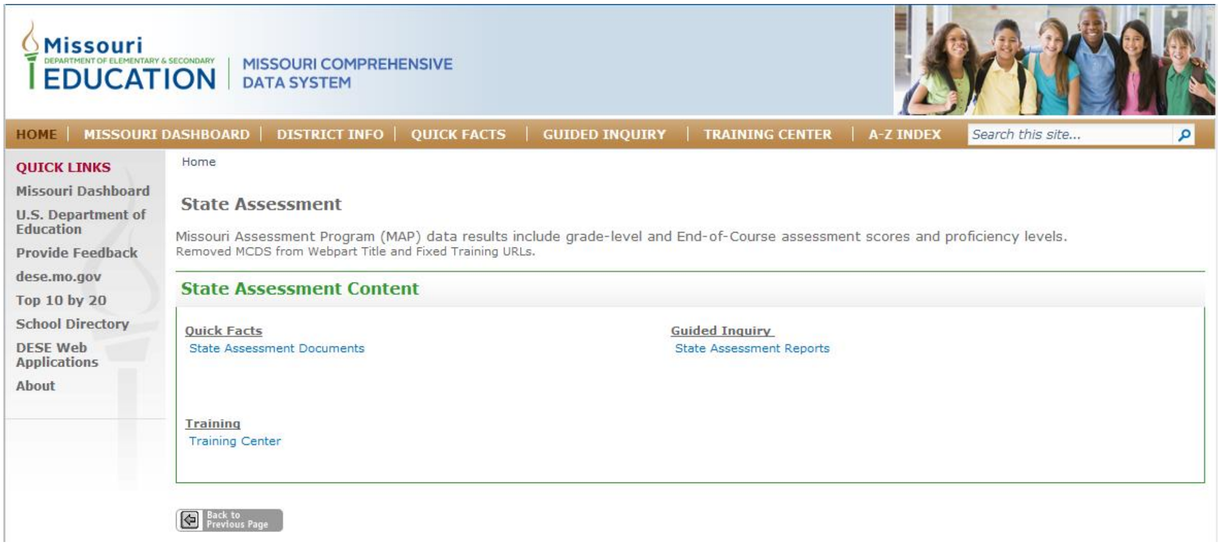
Beginning with the 2011–2012 school year, DESE transitioned all assessment reporting to the state's data portal, the Missouri Comprehensive Data System (MCDS). MCDS provides the general public with access to high-level EOC summary reports and allows school district personnel with appropriate permissions to access EOC data at a variety of levels. Through MCDS, designated district personnel are able to request on-demand, customized reports that are configured and disaggregated in ways that best meet their needs for such activities as evaluating programs, revising curriculum, and improving teaching and learning.

Users access MCDS from a link to the portal on the Department's homepage (http://dese.mo.gov/). From there, they access the data portal directly through the MCDS link, as shown in the following image.



Secure content is available through a link at the top of the MCDS portal's homepage. District users with appropriate permissions can log in to access data. Once users have logged in, they are returned to the MCDS portal page where they can locate EOC data through the State Assessment link.

On the State Assessment page, a Guided Inquiry link allows users to create summary administrative reports, achievement level reports, and historical AYP reports. Authenticated users can also download student-level data from the Guided Inquiry link.



An unlimited number of reports with any configuration may be created through MCDS. In addition to administrative reports, the MCDS portal also provides an unlimited configuration of summary reports, as shown in Table 8.1, that are beyond the scope of this technical report. Additional information and training pertaining to MCDS capabilities are available on DESE's website at http://mcds.dese.mo.gov/trainingcenter/Pages/default.aspx.

Table 8.1. Reports Available on the MCDS Portal

Report Type	Report
Administrative Reports	Guided Inquiry - State Assessment Administrative: MAP Scale Score Summary
	Guided Inquiry - State Assessment Administrative: MAP Student Demographics
	Guided Inquiry - State Assessment Administrative: MAP Participation Invalidation
	Guided Inquiry - State Assessment Administrative: MAP Student Achievement Level
	Guided Inquiry - State Assessment Administrative: EOC History Report
Achievement Level Reports	Guided Inquiry - State Assessment Achievement Level - 4 Levels: Achievement Level 4 Report
	Guided Inquiry - State Assessment Achievement Level - 4 Levels: Achievement Level 4 Charts
Content Standards Report	Guided Inquiry - State Assessment Content Standard - Item Analysis: Content Standard Summary
Item Analysis Expanded Reports	Guided Inquiry - State Assessment Content Standard - Item Analysis: Content Standard IBD
	Guided Inquiry - State Assessment Content Standard - Item Analysis: Goal Process IBD

8.2.4. Administrative Reports

These reports provide student-level test data. Based on only the MO EOC assessment results, four reports are generated: MO EOC Scale Score Summary, MO EOC Student Demographic, Student Achievement Level, and Student Report. Additionally, a historical report of the student's EOC participation is located within the administrative reports. The following list describes the contents of each administrative report:

- MO EOC Scale Score Summary: This report lists each student in the school or district along with his or her MOSIS ID, testing year, content area, grade level, MO EOC scale score, and achievement level.

- MO EOC Student Demographic: This report lists all students in the school or district along with their date of birth (DOB), content area, MOSIS ID, district ID, and relevant demographic information, including whether the student has been in the district for less than a year, whether the student has been in the building for less than a year, whether the student is limited English proficient (LEP), the student's race, whether the student qualifies for free and reduced lunch (FRL), whether the student has an individualized education program (IEP), whether the student is an English language learner (ELL)/LEP who has been in the school for less than one year and in the country for less than three years, whether the student is an LEP/ELL Title III, the number of months the LEP/ELL student has been in the United States, the student's disability diagnosis, and whether the student is Title I.

- Student Achievement Level: This report lists all students in a school or district along with the year of testing, content area, grade level, achievement level, and MOSIS ID.

- Student Report: For each school or district, this report contains the following information: student name, DOB, MOSIS ID, content area tested, grade level, achievement level, and scale score for each content area tested.

- EOC History Report: This report lists the history of MO EOC completion for all students in the school or district.

8.3. Summary of Test Score Results

The descriptive statistics for the number correct raw score, scale scores, and achievement levels for each of the seven MO EOC assessments from the Fall 2018 and Spring 2019 administrations are presented here. The statistics include *n*-counts, means, standard deviations (SD), minimum and maximum values, and a variety of data disaggregation.

8.3.1. Total Raw Scores

Table 8.2 summarizes the descriptive statistics for total raw score (RS) by test administration (test period) and content area. The information includes the total number of students who took the particular MO EOC Assessment (*n*-count), the number of items and possible points, the observed minimum and maximum scores, and mean and standard deviation of raw scores.

Table 8.2. Descriptive Statistics of Total Raw Scores

Test Period	Content Area	Core Form	*n*-Count	#Pts. Possible	Min.	Max.	Mean	SD
Fall 2018	English I	A	163	50	1	43	25.82	8.69
	English II	A	2,538	50	0	46	23.18	9.90
	Algebra I	A	5,307	50	0	50	21.26	10.70
	Algebra II	A	545	50	2	50	23.04	10.60
	Geometry	A	138	50	10	49	31.80	9.22
	Biology	A	1,315	50	0	50	26.58	11.40
		B	1,187	50	1	50	25.78	11.95
	Physical Science	A	39	50	10	35	20.82	6.75
Spring 2019	English I	C	5,794	50	0	47	27.81	8.18
		D	5,132	50	0	46	29.90	7.62
	English II	C	32,676	50	0	49	28.32	8.44
		D	27,410	50	0	48	29.26	7.92
	Algebra I	C	32,110	50	0	50	22.34	10.15
		D	26,514	50	0	48	22.56	9.04
	Algebra II	C	8,091	50	0	48	25.75	9.40
		D	7,520	50	0	48	24.05	9.12
	Geometry	C	1,839	50	4	48	21.42	9.01
		D	1,719	50	5	49	22.74	8.88
	Biology	A	33,918	50	0	50	29.73	10.17
		B	26,568	50	0	50	31.02	10.10
	Physical Science	A	2,358	50	5	49	27.13	8.80

8.3.1.1. <u>Total Raw Score by Cluster</u>

Tables 8.3 and 8.4 summarize the number correct RS—including the average raw score, the SD, and the standard error of measurement (SEM)—by test administration (test period), content area, and cluster. More information on SEM is provided in Chapter 7.

Table 8.3. Descriptive Statistics of Total Raw Scores by Cluster—Fall 2018

Content Area	Cluster	Core Form	#Pts. Possible	Mean	SD	SEM
English I	Reading Literary Texts	A	15	7.33	3.13	0.25
	Reading Informational Texts	A	15	7.91	3.20	0.25
	Writing	A	20	10.58	3.44	0.27
English II	Reading Literary Texts	A	15	6.09	3.01	0.06
	Reading Informational Texts	A	15	6.79	3.20	0.06
	Writing	A	20	10.30	4.79	0.10
Algebra I	Algebra	A	20	8.41	4.61	0.06
	Functions	A	18	7.40	4.00	0.05
	Number/Quantity and Statistics	A	12	5.44	3.11	0.04

Content Area	Cluster	Core Form	#Pts. Possible	Mean	SD	SEM
Algebra II	Algebra	A	27	13.19	6.38	0.27
	Functions	A	14	5.07	3.19	0.14
	Number/Quantity and Statistics	A	9	4.78	1.91	0.08
Geometry	Congruence/Similarity Coordinate Geometry & Circles	A	34	21.51	6.66	0.57
	Geometric Measurement & Modeling	A	10	6.34	2.16	0.18
	Statistics & Probability	A	6	3.96	1.26	0.11
Biology	From Molecules to Organisms: Structure and Process	A	13	5.76	2.97	0.08
	Ecosystems: Interactions, Energy, and Dynamics	A	10	4.96	2.89	0.08
	Heredity: Inheritance and Variation of Traits	A	11	5.63	2.80	0.08
	Biological Evolution: Unity and Diversity	A	12	7.76	3.13	0.09
	Earth and Human Activity	A	4	2.47	1.30	0.04
	From Molecules to Organisms: Structure and Process	B	12	5.68	2.77	0.08
	Ecosystems: Interactions, Energy, and Dynamics	B	12	6.02	3.54	0.10
	Heredity: Inheritance and Variation of Traits	B	12	6.27	3.26	0.09
	Biological Evolution: Unity and Diversity	B	11	6.34	2.80	0.08
	Earth and Human Activity	B	3	1.47	1.03	0.03
Physical Science	Matter and Its Interaction	A	14	4.28	2.16	0.35
	Motion and Stability: Forces and Interactions	A	14	7.18	2.74	0.44
	Energy	A	16	6.82	2.53	0.41
	Earth and the Universe	A	6	2.54	1.54	0.25

Table 8.4. Descriptive Statistics of Total Raw Scores by Cluster—Spring 2018

Content Area	Cluster	Core Form	#Pts. Possible	Mean	SD	SEM
English I	Reading Literary Texts	C	15	7.54	2.88	0.04
	Reading Informational Texts	C	15	7.53	2.79	0.04
	Writing	C	20	12.74	3.71	0.05
	Reading Literary Texts	D	15	8.34	2.97	0.04
	Reading Informational Texts	D	15	8.13	2.85	0.04
	Writing	D	20	13.44	3.05	0.04
English II	Reading Literary Texts	C	15	7.44	2.91	0.02
	Reading Informational Texts	C	15	7.98	3.04	0.02
	Writing	C	20	12.91	3.59	0.02
	Reading Literary Texts	D	15	7.78	2.94	0.02
	Reading Informational Texts	D	15	9.14	2.78	0.02
	Writing	D	20	12.35	3.50	0.02
Algebra I	Algebra	C	22	9.15	4.87	0.03
	Functions	C	19	8.46	4.26	0.02
	Number/Quantity and Statistics	C	9	4.73	1.98	0.01
	Algebra	D	18	9.35	3.85	0.02
	Functions	D	20	8.65	3.94	0.02
	Number/Quantity and Statistics	D	12	4.56	2.29	0.01
Algebra II	Algebra	C	27	13.36	5.86	0.07
	Functions	C	13	7.89	2.99	0.03
	Number/Quantity and Statistics	C	10	4.50	1.60	0.02
	Algebra	D	27	13.60	5.69	0.07
	Functions	D	13	6.08	2.70	0.03
	Number/Quantity and Statistics	D	10	4.37	1.78	0.02
Geometry	Congruence/Similarity, Coordinate Geometry, & Circles	C	35	15.59	7.14	0.17
	Geometric Measurement & Modeling	C	9	3.37	1.64	0.04
	Statistics & Probability	C	6	2.46	1.40	0.03
	Congruence/Similarity, Coordinate Geometry, & Circles	D	35	16.64	6.81	0.16
	Geometric Measurement & Modeling	D	7	2.83	1.59	0.04
	Statistics & Probability	D	8	3.27	1.64	0.04
Biology	From Molecules to Organisms: Structure and Process	A	13	6.49	2.71	0.01
	Ecosystems: Interactions, Energy, and Dynamics	A	10	5.66	2.70	0.01
	Heredity: Inheritance and Variation of Traits	A	11	6.33	2.69	0.01
	Biological Evolution: Unity and Diversity	A	12	8.61	2.69	0.01

Content Area	Cluster	Core Form	#Pts. Possible	Mean	SD	SEM
	Earth and Human Activity	A	4	2.63	1.25	0.01
Biology	From Molecules to Organisms: Structure and Process	B	12	6.96	2.45	0.02
	Ecosystems: Interactions, Energy, and Dynamics	B	12	7.38	3.12	0.02
	Heredity: Inheritance and Variation of Traits	B	12	7.70	2.90	0.02
	Biological Evolution: Unity and Diversity	B	11	7.22	2.36	0.01
	Earth and Human Activity	B	3	1.75	1.01	0.01
Physical Science	Matter and Its Interaction	A	14	6.57	2.95	0.06
	Motion and Stability: Forces and Interactions	A	14	8.58	2.98	0.06
	Energy	A	16	8.68	3.25	0.07
	Earth and the Universe	A	6	3.31	1.43	0.03

8.3.2. Scale Scores

Table 8.5 summarizes the descriptive statistics of scale scores for each MO EOC assessment by administration and core form. For the Fall 2018 and Spring 2019 administrations of English, Mathematics, and Science content areas, the lowest obtainable scale score is 325. The highest scale score has not been determined.

Table 8.5. Descriptive Statistics of the Scale Scores

Test Period	Content Area	Core Form	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	A	163	330	435	398.92	15.94
	English II	A	2538	325	444	394.00	17.00
	Algebra I	A	5307	329	471	396.10	15.31
	Algebra II	A	545	352	465	396.30	15.18
	Biology	A	1315	325	459	390.28	17.21
		B	1187	333	461	389.48	15.10
	Geometry	A	138	377	462	412.19	15.91
	Physical Science	A	39	365	407	384.74	10.75
Spring 2019	English I	C	5,969	325	456	401.40	15.33
		D	5,132	325	445	402.59	14.13
	English II	C	34,145	325	468	401.55	15.35
		D	27,410	325	458	403.04	14.02
	Algebra I	C	33,233	331	471	398.02	13.36
		D	26,524	331	471	399.14	12.28
	Algebra II	C	8,116	334	438	399.64	11.19
		D	7,520	335	443	399.73	10.98
	Geometry	C	1,857	366	452	398.83	12.82
		D	1,719	372	459	399.92	11.85

Test Period	Content Area	Core Form	*n*-Count	Min.	Max.	Mean	SD
	Biology	A	33,918	325	459	394.54	17.12
		B	26,568	325	461	396.62	14.79
	Physical Science	A	2,358	350	465	395.25	14.82

8.3.2.1. Scale Score by Demographic Group

Descriptive statistics of scale scores by demographic groups are summarized in Appendix F. The results are only reported for groups with 10 or more students. The demographic variables included are gender, ethnicity, migrant status, free and reduced lunch (FRL), limited English proficient (LEP), Title I, individualized education program (IEP), and accommodations.

8.3.3. Achievement Level Results

Questar monitors the performance level results using the preliminary and final data files for all new operational core forms as they enter. Careful monitoring of results will ensure unexpected findings are identified early. The performance level results for content areas with pre-equated forms, were examined for Fall 2018. The preliminary data and the final data results were reported to DESE. The results presented here are the final results for each administration. The achievement level distributions for the Fall 2018 and Spring 2019 administrations of English, Mathematics, and Science contents areas are presented by core form in the Tables 8.6 and 8.7.

Table 8.6. Achievement Level Distributions for Fall 2018

Content Area	Core Form	Statistics	Below Basic	Basic	Proficient	Advanced	Below Basic + Basic	Proficient + Advanced	Total
English I	A	Freq.	28	49	61	25	77	86	163
		%	17	30	37	15	47	53	100
English II	A	Freq.	798	717	876	146	1,515	1,022	2,537
		%	31	28	35	6	60	40	100
Algebra I	A	Freq.	1,803	1,473	893	1,137	3,276	2,030	5,306
		%	34	28	17	21	62	38	100
Algebra II	A	Freq.	163	126	177	79	289	256	545
		%	30	23	32	14	53	47	100
Geometry	A	Freq.	10	22	40	66	32	106	138
		%	7	16	29	48	23	77	100
Biology	A	Freq.	436	498	219	162	934	381	1,315
		%	33	38	17	12	71	29	100
	B	Freq.	402	434	212	139	836	351	1,187
		%	34	37	18	12	70	30	100
Physical Science	A	Freq.	18	18	3	0	36	3	39
		%	46	46	8	0	92	8	100

Table 8.7. Achievement Level Distributions for Spring 2019

Content Area	Core Form	Statistics	Below Basic	Basic	Proficient	Advanced	Below Basic + Basic	Proficient + Advanced	Total
English I	C	Freq.	764	1,713	2,378	1,113	2,477	3,491	5,968
		%	13	29	40	19	42	58	100
	D	Freq.	449	1,508	2,143	1,032	1,957	3,175	5,132
		%	9	29	42	20	38	62	100
English II	C	Freq.	4,669	9,687	16,121	3,667	14,356	19,788	34,144
		%	14	28	47	11	42	58	100
	D	Freq.	2,436	8,031	13,806	3,136	10,467	16,942	27,409
		%	9	29	50	11	38	62	100
Algebra I	C	Freq.	8,419	9,930	7,480	7,402	18,349	14,882	33,231
		%	25	30	23	22	55	45	100
	D	Freq.	5,570	8,413	6,692	5,847	13,983	12,539	26,522
		%	21	32	25	22	53	47	100
Algebra II	C	Freq.	1,058	3,082	2,522	1,452	4,140	3,974	8,114
		%	13	38	31	18	51	49	100
	D	Freq.	1,082	2,605	2,511	1,322	3,687	3,833	7,520
		%	14	35	33	18	49	51	100
Geometry	C	Freq.	308	745	561	243	1,053	804	1,857
		%	17	40	30	13	57	43	100
	D	Freq.	195	725	550	249	920	799	1,719
		%	11	42	32	14	54	46	100
Biology	A	Freq.	6,472	14,948	7,785	4,706	21,420	12,491	33,911
		%	19	44	23	14	63	37	100
	B	Freq.	3,680	11,517	7,018	4,345	15,197	11,363	26,560
		%	14	43	26	16	57	43	100
Physical Science	A	Freq.	455	1,030	656	217	1,485	873	2,358
		%	19	44	28	9	63	37	100

8.3.3.1. Achievement Level Distribution by Demographic Group

Achievement level distribution by demographic groups are summarized in Appendix F. The results are only reported for groups with 10 or more students. The demographic variables included are gender, ethnicity, migrant status, free and reduced lunch (FRL), limited English proficient (LEP), Title I, individualized education program (IEP), and accommodations.

References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. Washington, D.C.: AERA.

Baker, E. L., & Linn, R. L. (2002). *Validity issues for accountability systems*. Technical report 585. Los Angeles: Center for the Study of Evaluation.

Brennan, R. L. (2004a). BB-CLASS: Beta-Binomial Classification Consistency and Accuracy (Version 1.1) [Computer Software].

Brennan, R. L. (2004b). *Manual for BB-CLASS: A computer program that uses the beta-binomial model for classification consistency and accuracy*. Iowa City: Center for Advanced Studies in Measurement and Assessment (CASMA).

Chen, W.H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. Journal of Educational and Behavioral Statistics, *22*(3), 265–289.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

Dorans, N. J., & Holland, P. W. (1992). *DIF detection and description: Mantel-Haenszel and standardization* (ETS Research Report No. RR-92-10).

Fleiss, L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement, 33,* 613–619.

Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.

Huynh, H. & Meyer, J. (2010). Use of Robust z in Detecting Unstable Items in Item Response Theory Models. *Practical Assessment, Research and Evaluation, 15*(2).

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: Praeger. Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York: Springer-Verlag.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (1996). Standard setting: A bookmark approach. In D.R. Green (Chair), IRT based standard setting procedures utilizing behavior anchoring. Symposium conducted at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.

Linacre, J. M. (2015). Winsteps ® Rasch Measurement (Version 3.90.2) [Computer software]. Chicago, IL: Winsteps.com.

Linacre, J. M. (2006). *A user's guide to WINSTEPS Rasch-model computer programs*. Chicago: Winsteps.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32*, 179-197.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of National Cancer Institute, 22,* 719–748.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149–174.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives,* (pp. 249–281). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved December 21, 2009, from http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, *8*(2), 125-145.

Appendix A: Target and Actual Point Distributions in the Blueprints

English I

Table A.1. Actual Point Distributions—English I, Fall 2018

Reporting Category	Blueprint Target	Actual
	#Points	
Reading Literary Text	15	15
Reading Informational Text	15	15
Writing	20	20
Total	**50**	50

Table A.2. Actual Point Distributions—English I, Spring 2019

Core 1	Blueprint Target	Actual
Reporting Category	#Points	#Points
Reading Literary Text	15	15
Reading Informational Text	15	15
Writing	20	20
Total	**50**	50

Core 2	Blueprint Target	Actual
Reporting Category	#Points	#Points
Reading Literary Text	15	15
Reading Informational Text	15	15
Writing	20	20
Total	**50**	50

English II

Table A.3. Actual Point Distributions—English II, Fall 2018

Reporting Category	Blueprint Target	Actual
	#Points	
Reading Literary Text	15	15
Reading Informational Text	15	15
Writing	20	20
Total	**50**	50

Table A.4. Actual Point Distributions—English II, Spring 2019

Core 1 Reporting Category	Blueprint Target	Actual
	#Points	
Reading Literary Text	15	15
Reading Informational Text	15	15
Writing	20	20
Total	**50**	50

Core 2 Reporting Category	Blueprint Target	Actual
	#Points	
Reading Literary Text	15	15
Reading Informational Text	15	15
Writing	20	20
Total	**50**	50

Algebra I

Table A.5. Actual Point Distributions—Algebra I, Fall 2018

	Blueprint Target	Actual
Reporting Category	**#Points**	
Algebra	18-22	20
Function	18-22	18
Number and Data	8-12	12
Total	**50**	**50**

Table A.6. Actual Point Distributions—Algebra I, Spring 2019

Core 1	Blueprint Target	Actual
Reporting Category	**#Points**	
Algebra	18-22	22
Function	18-22	19
Number and Data	8-12	9
Total	**50**	**50**

Core 2	Blueprint Target	Actual
Reporting Category	**#Points**	
Algebra	18-22	18
Function	18-22	20
Number and Data	8-12	12
Total	**50**	**50**

Algebra II

Table A.7. Actual Point Distributions—Algebra II, Fall 2018

Reporting Category	Blueprint Target	Actual
	#Points	
Algebra	25-28	27
Function	11-14	14
Number Quantity and Statistics	10-12	9
Total	**50**	**50**

Table A.8. Actual Point Distributions—Algebra II, Spring 2019

Core 1	Blueprint Target	Actual
Reporting Category	#Points	
Algebra	25-28	27
Function	11-14	13
Number Quantity and Statistics	10-12	10
Total	**50**	**50**

Core 2	Blueprint Target	Actual
Reporting Category	#Points	
Algebra	25-28	27
Function	11-14	13
Number Quantity and Statistics	10-12	10
Total	**50**	**50**

Geometry

Table A.9. Actual Point Distributions—Geometry, Fall 2018

Reporting Category	Blueprint Target	Actual
	#Points	
Congruence/Similarity, Coordinate Geometry and Circles	32-35	34
Geometric Measurement and Modeling	6-10	10
Statistics and Probability	6-10	6
Total	**50**	**50**

Table A.10. Actual Point Distributions—Geometry, Spring 2019

Core 1	Blueprint Target	Actual
Reporting Category	#Points	
Congruence/Similarity, Coordinate Geometry and Circles	32-35	35
Geometric Measurement and Modeling	6-10	9
Statistics and Probability	6-10	6
Total	**50**	**50**

Core 2	Blueprint Target	Actual
Reporting Category	#Points	
Congruence/Similarity, Coordinate Geometry and Circles	32-35	35
Geometric Measurement and Modeling	6-10	7
Statistics and Probability	6-10	8
Total	**50**	**50**

Biology

Table A.11. Actual Point Distributions—Biology, Fall 2018 and Spring 2019

Core 1 Reporting Category	Blueprint Target #Points	Actual #Points
From Molecules to Organisms: Structure and Process	11-15	13
Ecosystems: Interactions, Energy, and Dynamics	8-12	10
Heredity: Inheritance and Variation of Traits	11-15	11
Biological Evolution: Unity and Diversity	11-15	12
Earth and Human Activity	3-6	4
Total	**50**	**50**

Core 2 Reporting Category	Blueprint Target #Points	Actual #Points
From Molecules to Organisms: Structure and Process	11-15	12
Ecosystems: Interactions, Energy, and Dynamics	8-12	12
Heredity: Inheritance and Variation of Traits	11-15	12
Biological Evolution: Unity and Diversity	11-15	11
Earth and Human Activity	3-6	3
Total	**50**	**50**

Physical Science,

Table A.12. Actual Point Distributions—Physical Science, Fall 2018 and Spring 2019

Core 1	Blueprint Target	Actual
Reporting Category	#Points	
Matter and Its Interactions	12-16	14
Motion and Stability: Forces and Interactions	12-16	14
Energy	12-16	16
Earth and the Universe	6-9	6
Total	**50**	**50**

Personal Finance

Table A.13. Actual Point Distributions—Personal Finance, Fall 2018

	Blueprint Target	Actual
Reporting Category	#Points	
I. Financial Decision Making/ II. Earning Income	10-12	13
III. Buying Goods and Services	10-12	7
IV. Savings/V. Using Credit	15-18	11
VI. Protecting and Insuring/ VII. Financial Investing	10-12	7
Total	**50**	**50**

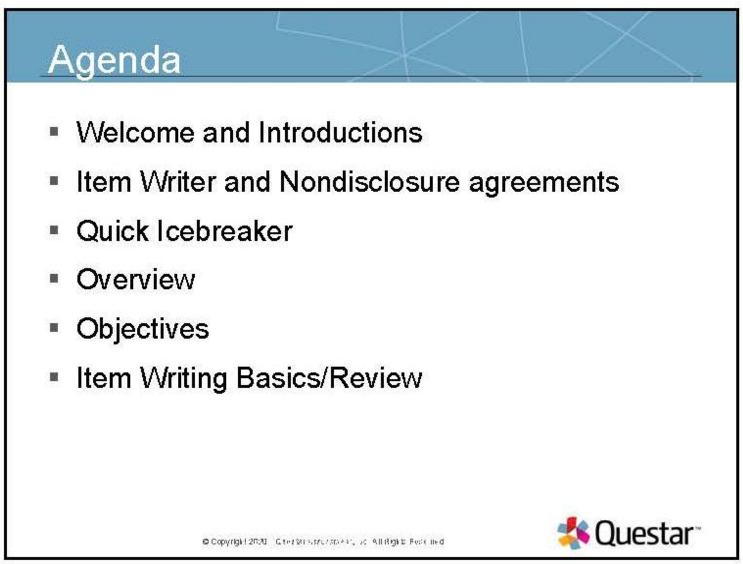
Table A.14. Actual Point Distributions—Personal Finance, Spring 2019

	Blueprint Target	Actual
Reporting Category	#Points	
I. Financial Decision Making/ II. Earning Income*	10-12	13
III. Buying Goods and Services	10-12	7
IV. Savings/V. Using Credit	15-18	11
VI. Protecting and Insuring/ VII. Financial Investing	10-12	7
Total	**50**	**50**

***Note:* This form and Fall 2018's form included two items that don't match to BP. Items are from an old test but mapped to new categories.

Appendix B: Item Writer Workshop Training Slides

1



2

Welcome and Introductions

MO Department of Elementary and Secondary Education

Lisa Sireno, Standards and Assessment Administrator

Shaun Bates, Director of Assessment

Debbie Jameson, Director of English Language Arts

Lisa Scroggs, Assistant Director of English Language Arts

Kristen McKinney, Director of Science

Dixie Grupe, Director of Social Studies

Chip Sharp, Director of Mathematics

Questar

3

Welcome and Introductions

Questar Support

Adam Johnson, Sr. Program Manager

Steven Daniels, Associate Program Manager

Jennifer Wright, Assessment Development Manager

Les Sewall, Assessment Development Manager

Vince Thomas, Field Systems Engineer

Questar

4

Welcome and Introductions

Questar Content Leads

Nancy McDonald, ELA

Michelle Udvare, Science

Tim Sitar, Social Studies

Olga Garza, Mathematics

Questar

5

Welcome and Introductions

Questar Facilitators

Bill Gleason, English I

Nancy McDonald, English II

Wendi Patrick, Government

John Haglund, American History

Jean Sofia, Algebra I

John Gunning, Algebra II

Joyce Jonik, Geometry

Stephanie Ryan, Physical Science

Stephanie Shaw, Biology

Questar

6

Housekeeping

- Forms
  - Non-disclosure Agreements
  - Reimbursement and Stipend Forms
- Breakfast, lunch, and breaks
- Restrooms

Questar

7

Schedule – Day 1

Time		Activity
7:00 - 8:00	AM	Registration/Breakfast
8:00 - 10:00	AM	Introductions, Review Agenda, Training
10:00 - 12:00	AM	Item Writing Breakout Sessions
12:00 - 1:00	PM	Lunch
1:00 - 3:00	PM	Continue Item Writing
3:00 - 3:15	PM	Afternoon Break
3:15 - 4:30	PM	Continue Item Writing

Questar

8

Schedule – Days 2 - 4

Time		Activity
7:00 - 8:00	AM	Breakfast
8:00 - 10:30	AM	Continue Item Writing
10:30 - 10:45	AM	Morning Break
10:45 - 12:00	PM	Continue Item Writing
12:00 - 1:00	PM	Lunch
1:00 - 3:00	PM	Continue Item Writing
3:00 - 3:15	PM	Afternoon Break
3:15 - 4:30	PM	Continue Item Writing

Questar

9

How to write good …

1. Avoid alliteration. Always.
2. Prepositions are not words to end sentences with.
3. Avoid clichés like the plague. (They're old hat).
4. Eschew ampersands & abbreviations, etc.
5. One should never generalize.
6. Comparisons are as bad as clichés.
7. Be more or less specific.
8. Sentence fragments? Eliminate.

Nine. Be consistent.

10. Who needs rhetorical questions?

Questar

10

MO EOC Assessment Overview

- Part of the Missouri school and accountability system
- Criterion referenced
- Designed to assess Missouri Learning Standards
- Founded on educator involvement and feedback
- Starts with items authored by Missouri educators

Questar

11

Assessment Overview

- **Assessments include**
  – English I, English II
  – Algebra I, Algebra II, Geometry
  – Biology, Physical Science
  – American History, Government

Questar

12

Objectives

Develop an understanding of best practices in item development

Write items that meet Missouri Learning Standards and are grounded in industry best practices

Questar

13

Purpose

To write test items that

- Measure what they claim to measure
- Align to a standard/skill
- Are clear to the student
- Are accessible to all students
- Mirror the classroom experience

Questar

14

Test Security and Integrity

- The importance of this cannot be overstated.
- You will sign an NDA.

15

Test Security and Integrity

- Test questions must remain secure.
- All test questions and passages, whether draft or final, are to be regarded as secure.
- Materials may not be reproduced, discussed, or in any way released or distributed. This includes emailing, copying, printing, posting or taking screenshots.

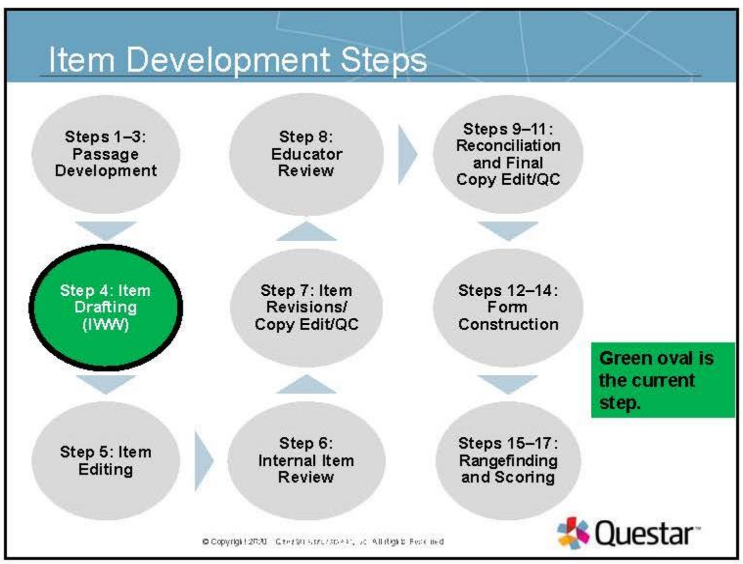
16

Test Security and Integrity

- Phones are put away and on silent or turned off.
- Use only the materials provided and leave all secure materials in the room. They will be collected and destroyed at the conclusion of the meeting.
- Refrain from using the internet for personal reasons or accessing social media in the meeting room.

Questar

17



18

Universal Design - Review

- Eliminating or minimizing barriers
- Accessible for all students

Questar

19

Bias and Sensitivity - Review

Construct-irrelevant factors may influence student performance causing unfair advantage or disadvantage to any group of students.

Questar

20

Pop Quiz: Terms to Know

- Item
- Item Stem
- Passage/Stimulus
- Rubric
- Sample Answer/Solution

Questar

21

Terms to Know

- Rationale
- Standard Code
- Tools/Practice Code
- Depth of Knowledge (DOK)

Questar

22

Item Writing Basics Review

Describe a basic principle or guideline for good item writing.

23

Style and Format

- Use and format emphasis words
  - Best, mainly, most likely, main
- Use direction lines as appropriate
  - "Select **all** that apply" for multi-select
- Keep art as clear as possible
  - Detailed, but not too busy
  - Clear and complete descriptions or insert concept art
- Online presentation
  - Scrolling considerations
  - Option format: vertical, horizontal, stacked

24

Group Exercise: revise this item

ACME was founded in

A. 1958.

B. 1959.

C. 1960.

D. 1962.

25

Group Exercise: revise this item

Which of the following products does ACME produce and when did production begin?

A. Irons, 1960

B. Waffle makers, 1961

C. Can openers, 1967

D. Toasters, 1968

26

Group Exercise: revise this item

Renee Zellweger stars as a show girl in Chicago in which of the following films?

A. Chicago
B. Music Man
C. Wizard of Oz
D. Sound of Music

27

DOK Review

Level 1 involves…

Level 2 involves…

Level 3 involves…

Level 4 involves…

28

Breakout Sessions Agenda

- Facilitators will oversee and coordinate activities
- Group introductions and overview of activities
- Chrome Books for each participant
- Distribution and overview of materials
- Start authoring items in templates

Questar

29

Breakout Rooms by Content Area

Content Area	Room
English I	**Escollo**
English II	**Lookout**
Government	**Palma**
American History	**Valderrama**
Algebra I	**Terrace II**
Algebra II	**Terrace I**
Geometry	**Board Room**
Physical Science	**Madrid**
Biology	**Barcelona**

Questar

30

Next Steps

- Let's take a quick break
- Reconvene in individual breakout rooms

Questar

31

Appendix C: Performance Level Setting Report

Missouri
DEPARTMENT OF ELEMENTARY & SECONDARY
EDUCATION

Missouri Assessment Program
Grades 5 and 8 Science, Physical Science, and Biology

Performance Level Setting 2019 Draft Technical Report

Prepared for the
Department of Elementary & Secondary Education

Data Recognition Corporation
Maple Grove, MN 55311

ACS Ventures
Las Vegas, NV 89135

DATA RECOGNITION DRC CORPORATION

ACS VENTURES BRIDGING THEORY & PRACTICE

Table of Contents

A
Executive Summary

Executive Summary

Summary

- On July 16–18, 2019, the Missouri Department of Elementary and Secondary Education (DESE) partnered with Data Recognition Corporation (DRC), Questar Assessment Inc. (QAI) and ACS Ventures to conduct a performance level setting, commonly referred to as a standard setting, for the Missouri Assessment Program (MAP) tests of grades 5 and 8 science, Physical Science, and Biology.
- The *standard setting* was needed because of implementation of new Missouri Learning Standards in 2016. The MAP assessed these standards for the first time in spring 2019.

Background

In 2016, the Missouri State Board of Education approved new Missouri Learning Standards for science, and these standards were implemented in 2017–18. The MAP began assessing these standards in 2018–19. As part of a multi-phased standard setting, DESE sought to establish new cut points for MAP science assessments which: (a) reflect the new Missouri Learning Standards, (b) link students' scores on the MAP to the state's expectations for students in each performance level, and (c) are well articulated across grades and courses.

The performance levels for MAP are designed to indicate students' knowledge of the skills listed in the Missouri Learning Standards. The performance levels are *Below Basic, Basic, Proficient*, and *Advanced*.

Standard Setting Methodology

A total of 47 Missouri educators engaged in the Bookmark Standard Setting Procedure (Lewis, Mitzel, & Green, 1996; Lewis, Mitzel, Mercado, & Schulz, 2012) to recommend cut scores. This method has been used on assessments in Missouri and across the nation, including for English language arts and mathematics of the MAP. There were 13 participants for the grade 5 assessment, 11 for grade 8, 12 for Biology, and 11 for Physical Science.

Participants studied the updated Missouri performance level descriptors (PLDs) and Missouri Learning Standards to review the knowledge, skills, and abilities expected of students in each performance level. Each performance level was associated with a level of mastery of the Missouri Learning Standards. Participants then discussed the content-based expectations for students at the threshold of each performance level (e.g., a student who is just *Proficient*).

Participants studied *ordered item booklets* (OIBs) that comprised collections of operational test items that were ordered by difficulty. A separate OIB was created for each test, and items' difficulty values were based on students' performance on the test items. Participants studied the OIBs to understand the knowledge and skills measured by the tests.

On May 2nd, a benchmark panel of Missouri educators and stakeholders was convened to help DESE select appropriate benchmarks for the four science tests. This panel also provided contextual information to accompany the benchmarks when they were ultimately used at the performance level setting workshop. For grades 5 and 8 science, *benchmarks* based on NAEP were presented for participants' consideration as they recommended their *Proficient* cut scores. The benchmarked cut scores, when applied to Missouri students' scores, categorized approximately the same percentages of students as *Proficient* or above on MAP and on NAEP. A band of ±1 conditional standard error of measurement (CSEM) was used to create a band that was reflected in participants' OIBs. Participants were told that if they recommended cut scores in this range, the percentage of students classified as *Proficient* and above on MAP would be similar to that on NAEP. Participants were encouraged to focus their attention on this range in the OIB, and that it was likely (but not compulsory) that their *Proficient* cut scores be within this range.

Unlike the grade 5 and grade 8 assessments, a parallel NAEP assessment program does not exist for the EOC assessments. NAEP has a grade 12 assessment, but it does not provide statewide results and not all states participate in the administration of the assessment. At the benchmark panel meeting held on May 2nd, the panel endorsed the use of NAEP for the grade 5 and grade 8 assessments, but it could not find suitable external benchmarks for the EOC assessments. The benchmark panel developed recommendations for benchmarks for the two EOC assessments, but it did not express a significant amount of confidence in the recommendations. Just as with grade 5 and 8, a band of ±1 CSEM was used to create a benchmark band, loosely associated with a *Proficient range*, and this range was reflected in participants' OIBs. However, participants were told that the information was provided simply for the participants' information, and that the participants' recommendation would not necessarily be consistent with the benchmark band.

Participants engaged in three rounds of individual judgments and group discussions. In each round, participants recommended cut scores by considering the content-based expectations for students in each performance level, and then identifying the sets of items in their OIBs which best represented these expectations. By *placing bookmarks*, participants recommended cut scores on the test scale.

Between rounds, participants were shown feedback (e.g., median bookmarks, impact data). The committees' median judgments were taken as their recommendations. For rounds 1 and 2, feedback was provided to each individual panel and only shown the results for their specific program. After round 3 ratings were completed, all participants were shown the results across all grade levels and programs to allow the entire panel to discuss the results and evaluate the consistency of their specific program with the other Science assessments.

After the round 3 discussion, table leaders convened to examine the recommendations. As needed, table leaders recommended adjustments to promote articulation among the performance standards across grades. When examining the cut scores, the table leaders recommended adjustments to two cut scores to improve the articulation of the entire system of performance standards. The table leaders noted that the percent of students classified as *Below Basic* in grade 8 was markedly lower than observed in other grades, so it recommended adjusting its recommended bookmark to 17 from 15. Similarly, the table leaders reported that the percentage of Biology students classified as *Advanced* was unexpectedly high, especially when compared to the other grades, and recommended adjusting their bookmark to 84 from 83.

Several table leaders noted that the percentage of students classified as *Proficient* and above in grade 8 was markedly higher than that observed in the other three grades. These table leaders noted that this pattern was unexpected; however, the table leaders could not make a consensus recommendation to adjust the cut score. After the standard setting, this cut score was adjusted by one conditional standard error of measurement (CSEM) value: the CSEM quantifies the amount of statistical noise around a point on the test scale. Accordingly, the grade 8 *Proficient* cut score was adjusted to 510 from 501, a change of +1 CSEM. This adjustment was made to promote articulation across grades while still honoring the judgments of Missouri educators at the standard setting.

Table 1 shows the recommended cut scores for the MAP science tests, plus the associated impact data using spring 2019 administration data. It should be noted that the cut score recommendations are recorded on the scale score metric for the grade 5 and grade 8 assessments, and on the theta level scale for the EOC assessments. Impact data are the percentages of students who would be classified in each performance level if the cut scores were applied to students' scores.

Table 1. Recommended Cut Scores and Associated Impact Data for the MAP Science Tests

		Recommended Cut Scores			% Students by Level Based on Spring 2019				
Content	**Grade**	***Basic***	***Proficient***	***Advanced***	***Below Basic***	***Basic***	***Proficient***	***Advanced***	***Prof. + Adv.***
Science	**5**	275	310	344	26.0%	31.1%	29.3%	13.5%	42.8%
	8	468	510	537	20.4%	35.7%	25.9%	18.1%	43.9%
	PS	382	400	417	19.2%	42.9%	28.2%	9.7%	37.9%
	BIO	381	400	411	16.8%	43.8%	24.5%	15.0%	39.5%

Statistical Standard Error Estimates

For each recommended cut point, the conditional standard error of measurement (CSEM) was calculated at that point on the operational test form. Using these CSEM values, the corresponding cut score plus or minus 1 and 2 CSEM intervals was calculated. The resulting cut scores and the corresponding impact data are provided in Table 2.

Table 2. Cut Scores Adjusted by CSEM Intervals and Associated Impact Data for the MAP Science Tests

		Cut Scores Adjusted by CSEM				Associated Impact Data			
		5	**8**	**PS**	**BIO**	**5**	**8**	**PS**	**BIO**
+2 CSEM	Below Basic	--	--	--	--	44.45%	36.43%	40.80%	32.58%
	Basic	297	490	392	389	33.25%	37.82%	41.26%	48.99%
	Proficient	332	528	410	409	18.80%	18.99%	16.41%	14.27%
	Advanced	368	555	429	421	3.50%	6.76%	1.53%	4.17%
	Prof + Adv					*22.30%*	*25.75%*	*17.94%*	*18.44%*
+1 CSEM	Below Basic	--	--	--	--	34.53%	27.66%	28.84%	24.96%
	Basic	286	479	387	385	33.27%	37.79%	42.88%	46.14%
	Proficient	321	519	405	404	24.86%	22.98%	24.64%	19.97%
	Advanced	356	546	423	416	7.34%	11.57%	3.65%	8.93%
	Prof + Adv					*32.20%*	*34.55%*	*28.29%*	*28.90%*
As Recom-mended	Below Basic	--	--	--	--	26.05%	20.38%	19.17%	16.77%
	Basic	275	468	382	381	31.11%	35.71%	42.92%	43.77%
	Proficient	310	510	400	400	29.32%	25.86%	28.20%	24.49%
	Advanced	344	537	417	411	13.53%	18.05%	9.71%	14.97%
	Prof + Adv					*42.85%*	*43.91%*	*37.91%*	*39.46%*
-1 CSEM	Below Basic	--	--	--	--	18.82%	14.54%	10.01%	10.79%
	Basic	264	457	377	377	27.53%	32.37%	37.74%	39.73%
	Proficient	299	501	395	396	31.35%	27.34%	34.31%	25.99%
	Advanced	332	528	410	406	22.30%	25.75%	17.94%	23.50%
	Prof + Adv					*53.65%*	*53.09%*	*52.25%*	*49.49%*
-2 CSEM	Below Basic	--	--	--	--	12.91%	9.99%	3.44%	5.75%
	Basic	253	446	371	373	23.44%	28.21%	33.29%	32.53%
	Proficient	288	492	390	391	30.45%	27.25%	31.89%	25.72%
	Advanced	320	519	404	402	33.20%	34.55%	31.38%	35.99%
	Prof + Adv					*63.65%*	*61.80%*	*63.28%*	*61.71%*

B

Standard Setting Methodology and Recommendations

Performance Level Setting Methodology

On July 16–18, 2019, the Missouri Department of Elementary and Secondary Education (DESE) sponsored a performance level setting[1] for the Missouri Assessment Program (MAP) tests of grade 5 science, grade 8 science, end-of-course (EOC) Physical Science, and EOC Biology. To conduct the performance level setting, DESE partnered with Data Recognition Corporation (DRC), Questar Assessment Inc. (QAI), and ACS Ventures (ACS). DRC, ACS, and QAI facilitated the performance level setting on behalf of DESE.

The performance level setting was needed because of the implementation of the Missouri Learning Standards (MLS), the state's new set of academic content standards. Four performance levels are designed to indicate students' knowledge of the skills listed in the MLS. These performance levels, based on students' performance on the MAP assessments, are *Below Basic, Basic, Proficient*, and *Advanced*.

About this Section

This section details the planning of the performance level setting, the implementation of the workshop, the analysis of Missouri educators' recommendations, and the consideration of the cut scores by DESE. A summary of this work can be found in Section A of this report. Further details about the workshop, such as workshop agendas and detailed presentations of participants' recommendations, can be found in subsequent sections of this report.

Background

In 2005, cut scores were established for the MAP that linked the performance standards to state NAEP performance. This explicit link was codified state law (i.e., Missouri SB 1080); however, this law is no longer in effect. DESE has indicated that NAEP-like performance standards have benefitted the state, and this linkage to NAEP served the state well.

Missouri later joined the Smarter Balanced Assessment Consortium (SBAC). The State left SBAC in the summer of 2015, and Missouri administered a new assessment in 2015–16. New performance standards, including performance level descriptors (PLDs) and cut scores were established in July 2016.

In 2016, the Missouri State Board of Education approved the new Missouri Learning Standards (MLS). For English language arts (ELA) and mathematics, the MLS were implemented in Missouri schools in school year 2016–17, and the MAP began assessing them in school year 2017–18. For science, the MLS were implemented in schools in school year 2017–18, and the MAP assessed them from 2018–19.

In July 2018, DESE sponsored a performance level setting for ELA and mathematics that reflected the new MLS. To continue this transition to the new MLS, DESE sponsored a performance level setting for the science assessments, using a similar methodology, in July 2019. DESE sought to establish new cut

[1] The literature commonly refers to a *performance level setting* as a *standard setting*. In this context, both terms are synonymous. However, to prevent confusion with the process used to establish the state academic content standards (i.e., the Missouri Learning Standards), this section refers to the process used to establish performance standards as a *performance level setting*. At the workshop, the two terms were used interchangeably.

scores for MAP science assessments which achieve three objectives. Specifically, the performance standards were designed to:

(a) reflect the new MLS,

(b) link students' MAP scores to DESE's expectations for students in each performance level, and

(c) reflect appropriate rigor and consistency across grades and courses.

To achieve these goals, DESE partnered with DRC, QAI, and ACS to implement a performance level setting to recommend cut scores for the four MAP tests of science: grade 5 science, grade 8 science, Physical Science, and Biology.

Performance Level Setting Methodology and Rationale

The 2019 performance level setting for the MAP science assessments comprised three interconnected phases, shown here.

1) **Developing the performance level descriptors (PLDs).** On June 7–8, 2016, DESE worked with a committee of 50 Missouri educators to develop PLDs for the MAP science assessments. This work is summarized later in this section, and further information is presented in the PLD report (Buckendahl & Wiley, 2016).
2) **Defining the benchmarks for the performance level setting.** On May 2, 2019, DESE convened a committee of 12 Missouri educators and stakeholders to recommend benchmarks for the *Proficient* cut scores of the MAP science tests. The process used to select the benchmarks is presented later in this section under the heading "Benchmarks." Further information is found in Section J of this report.
3) **Performance level setting.** On July 16–18, 2019, DESE partnered with DRC, QAI and ACS to recommend performance standards for the MAP science tests. This process is described in this section.

The Bookmark Standard Setting Procedure (Lewis, Mitzel, & Green, 1996; Lewis, Mitzel, Mercado, & Schulz, 2012) was implemented to recommend cut scores for the MAP science tests. This method has been used on assessments in Missouri and across the nation (Karantonis & Sireci, 2006), including for the previous version of the MAP (Data Recognition Corporation, 2016) and the current version of MAP for English language arts and mathematics (Data Recognition Corporation, 2018).

As an item-mapping process, the Bookmark Procedure is particularly useful for large-scale assessments that include both multiple-choice and multi-point constructed-response items, like the MAP. Because Bookmark allows these different item types to be ordered together in *ordered item booklets*, and because of its history of use in Missouri and across the nation, DESE selected the Bookmark Procedure for the 2019 MAP science performance level setting.

The performance level setting also incorporated elements of the *evidence-based standard setting* framework (McClarty, Way, Porter, Beimers, & Miles, 2013). In particular, focused attention was paid before the performance level setting workshop to the types of performance standards that DESE would consider reasonable for MAP. Selected policy information was provided to performance level setting participants in the form of benchmarks (see Phillips, 2012), allowing educators to consider this policy information in an actionable way as they made their content-based judgments in the Bookmark Procedure.

Throughout the process to design and implement the performance level setting, DESE sought advice from its technical advisory committee (TAC). The TAC is composed of nationally-recognized experts in educational assessment. The TAC helped guide DESE in its choice of performance level setting methodology.

Performance Level Descriptors (PLDs)

PLDs summarize the knowledge, skills, and abilities expected of students in each performance level. On June 7–8, 2016, DESE partnered with ACS to develop performance level descriptors (PLDs) for the MAP assessments. A total of 50 Missouri educators took part in this effort for the MAP program; of these participants, eight educators focused on the PLDs for science.

The PLDs were developed using the MLS for science, thereby linking the content of the MLS with the expectations for students in each performance level. (Additional information about the PLD development process may be found in the PLD report; Buckendahl & Wiley, 2016). Egan, Schneider, and Ferrara (2012) suggest a framework of four types of PLDs, described here.

Policy PLDs summarize the DESE's definition for each performance level, providing information to stakeholders on the state's suggested interpretation of each level. They are typically not specific to any given grade or content area. The policy PLDs for the MAP (Missouri Department of Elementary and Secondary Education, 2017) are shown here, with emphasis added to show differences:

- *Below Basic.* Students performing at the *Below Basic* level on the Missouri Assessment Program demonstrate a <u>minimal command</u> of the skills and processes identified in the Missouri Learning Standards.
- *Basic.* Students performing at the *Basic* level on the Missouri Assessment Program demonstrate a <u>partial or uneven command</u> of the skills and processes identified in the Missouri Learning Standards.
- *Proficient.* Students performing at the *Proficient* level on the Missouri Assessment Program demonstrate an <u>adequate command</u> of the skills and processes identified in the Missouri Learning Standards.
- *Advanced.* Students performing at the *Advanced* level on the Missouri Assessment Program consistently demonstrate a <u>thorough command</u> of the skills and processes identified in the Missouri Learning Standards.

Range PLDs summarize the knowledge, skills, and abilities expected of students in a given performance level on a specific test. The range PLDs show the types of content, as informed by the state content standards, that should be mastered by students in each performance level on the test at hand. The range PLDs generally show these expectations for students across the range of performance for the performance level: for example, the *Proficient* PLD for a test summarizes skills held by students who are just barely in the *Proficient* level and also skills held by students who are nearly *Advanced*. Range PLDs are often shared with teachers and schools to help them understand the level of construct mastery expected of students in each performance level on each test.

Threshold PLDs are based on the range PLDs. They summarize the knowledge, skills, and abilities expected of students who are at the point-of-entry (the *threshold*) of each performance level. For any given test, these descriptors show the types of skills needed just to be classified in a given level (e.g., just to be classified in the *Proficient* level). These PLDs specify the content expectations for students with performance analogous to the cut points. These descriptors are typically used by participants at performance level setting workshops to help inform decisions they make about cut points.

Reporting PLDs are the version of the PLDs used for score reporting. Typically, the reporting PLDs comprise a version of the policy or range PLDs, and the language in the reporting PLDs is adjusted to be accessible to a wide audience that may not have in-depth content knowledge.

After the PLD workshop, DESE reviewed the PLDs for science. In early 2019, DESE reviewed the PLDs once again to determine whether the PLDs still reflected the knowledge, skills, and abilities that the state expected of students in each performance level. DESE determined that the PLDs written in 2016 did not always reflect the full range of knowledge, skills, and abilities expected of students in each performance level (i.e., did not always reflect the full *range* of performance within each level).

DESE worked with its internal content experts to refine the PLDs created by Missouri educators in 2016. These PLDs comprised policy PLDs and range PLDs, and they were provided to participants at the performance level setting to inform their judgments. At the performance level setting, participants discussed these PLDs, and participants developed informal threshold PLDs to inform their cut score recommendations. Reporting PLDs were not within the scope of the performance level setting.

Benchmarks

Benchmarks comprised an important component of the performance level setting process. Benchmarks refer to any external content- or policy-based information that is presented to participants that help participants make their cut point recommendations. The use of benchmarks at performance level setting is well established (Phillips, 2012; McClarty et al., 2013), especially in the Bookmark Procedure (Lewis, Mitzel, Mercado, & Schulz, 2012). Many states have used benchmarks to provide actionable, policy-based information to performance level setting participants. Participants can then bring their content-based expertise to bear, joining it with the benchmarks. Thoughtful use of benchmarks can bring policy- and content-based information together in a meaningful way.

To avoid overloading the performance level setting participants with information, the DESE's technical advisory committee (TAC) recommended that a limited number of benchmark sources be presented at the performance level setting (i.e., up to two), and that benchmarks be presented only for the *Proficient* cut point (i.e., not for the *Basic* or *Advanced* cut points). DESE accepted this recommendation, as the Department wanted Missouri educators to focus mostly on the content-based expectations of students in each performance level during the performance level setting. This practice was also used during the 2018 performance level settings for ELA and mathematics (Data Recognition Corporation, 2018).

Benchmark Committee

To recommend benchmarks for the performance level setting, DESE convened a committee of 12 Missouri educators and stakeholders on May 2, 2019. During this one-day meeting, the *benchmark* *committee* reviewed information from well-respected measures of student performance with the goals of (a) recommending benchmarks for use at the 2019 science performance level setting, and (b) recommending additional contextual information that would help participants at the performance level setting make the best use of the benchmarks.

For grades 5 and 8 science, DESE determined that benchmarks based on the National Assessment of Educational Progress (NAEP) would comprise suitable benchmarks for the performance level setting. During the meeting of the benchmark committee, DESE asked what additional, contextual information would help performance level setting participants make use of these benchmarks. The benchmark committee noted that NAEP tests science differently than MAP: NAEP includes an experiential component, NAEP reports *science practices* separately, and NAEP tests in grades 4 and 8 (instead of grades 5 and 8, as does MAP). Moreover, the committee noted that NAEP's cut scores were established by a different organization for a markedly different purpose. Accordingly, the benchmark committee suggested that (a) performance level setting participants be told of the differences between NAEP and MAP; (b) achievement level descriptors (ALDs) from NAEP be provided at the performance level setting, so participants could see the similarities and differences between the two sets of performance standards; and (c) participants at the performance level setting not be limited to only recommending cut scores consistent with the benchmark, as there were important content-based factors that may lead to *Proficient* cut scores outside the benchmark ranges.

DESE accepted these recommendations for grades 5 and 8 science. Benchmarks for these tests were based on the percentage of students classified as *Proficient* or above on NAEP Science in 2015, the last year for which data were available. Cut scores on the MAP were identified which yielded approximately equal percentages of students classified as *Proficient* and above. Acknowledging that NAEP and MAP are different tests, each benchmark was then transformed into a *benchmark range* (sometimes termed at the workshop as the *Proficient range*) of ±1 SEM. At the workshop, the benchmark range was shown to participants as a range of pages in the ordered item booklet (OIB). Participants were told that if they placed their *Proficient* bookmark within this range, their recommended *Proficient* cut score would be consistent with the benchmark. However, participants were also told this was not compulsory, and they should use the content-based expectations for students in each performance level as their primary decision-making factor when placing their bookmarks.

During the May benchmark panel meeting, the committee also discussed the availability of external benchmarks for the two EOC assessments. However, the identification of appropriate benchmarks was more challenging due to the fact that NAEP does not have a comparable assessment for either of the two EOC assessments. The content of the NAEP grade 12 assessment is not consistent with either test and the test taking population (i.e. the grade level of participants) is also not comparable. The scenario is further complicated by the fact that there is not a plethora of data on the Missouri test taking populations for the Biology and Physical Sciences exams.

The benchmark panel was able to identify benchmark ranges, rather than a single benchmark value. The committee felt that the use of ranges rather than a single point was indicative of the lack of firm external benchmarks that they could identify and support. For Biology, the committee stated that a benchmark range of 45 to 50 percent of students considered to be *Proficient* felt like the most appropriate range. For Physical Sciences, the committee stated that a benchmark range of 25 to 35 percent of students considered to be *Proficient* felt like the most appropriate range.

However, even with the expanded range representing the benchmark, the committee still expressed significant concerns regarding the use of the external benchmarks. As a result of these concerns, the performance level setting was still provided the range of items associated with the *Proficient range*. But panelists were also told about the concerns and issues raised by the benchmark panel. They were also told that there was not an explicit expectation that their eventual ratings would be consistent with the benchmarks identified. Instead, the information was provided to give panelists some additional context for their rating, but only for that purpose.

Additional information about the benchmark committee can be found in Section J of this report.

Performance Level Setting Committee Composition

The performance level setting committee gathered educators and stakeholders from across the state of Missouri. Participants for the committee were recruited for the workshop by DESE and were invited to the workshop by DRC and QAI. DESE took special care to invite workshop participants that:

a) were well qualified (e.g., had experience teaching in Missouri classrooms),
b) were diverse in terms of demographic characteristics (e.g., gender, ethnicity),
c) were diverse in terms of geographic location within the state, and
d) had knowledge of the tested content and population.

DESE used a matrix-based strategy in its recruitment efforts. Specifically, DESE sought to recruit a diverse set of educators to each of the four groups, purposefully including educators from across the state in each group.

The performance level setting committee comprised 47 Missouri educators and stakeholders. Of these participants, 45 completed a demographic questionnaire administered at the end of the workshop. The workshop evaluation, along with detailed results from the evaluation, are presented in Section I of this report. Selected findings from the demographic questionnaire are shown here.

- 69% of committee members were female, 29% were male, and 2% did not respond
- 93% were Caucasian, 2% were black, 2% were of two or more races, and 2% did not respond
- 60% were classroom teachers, and 20% were non-teacher educators
- 60% had taught for more than 10 years, and 30% had taught for more than 20 years
- 80% had a master's degree or higher
- 62% reported more than half of students at their school qualify for free or reduced-price meals
- 53% or more had experience teaching students in special education
- 22% had experience teaching English language learners

As shown here, most of the committee members were female and Caucasian. Most participants were classroom teachers, and approximately one-fifth were non-teacher educators such as content coaches. Four-fifths of committee members had a master's degree or higher, and more than half had taught for 10 years or more. Moreover, most participants had experience teaching students in special populations, such as students receiving special education services.

Performance Level Setting Committee Configuration

Educators were divided into groups by test: approximately 12 participants formed each group, and each group focused on one test. As such, there were four groups at the workshop: one each for grade 5 science, grade 8 science, Physical Science, and Biology. Each group worked in a separate room.

In each group, DESE assigned participants to two tables of approximately six participants each. Each group and each table was as representative as possible in terms of relevant demographics (e.g., gender, geographic location). In addition, DESE designated one participant at each table as the table leader. Table leaders were selected based on their ability to facilitate small-group discussions at their table; and their ability to represent their table's recommendations during the across-grade discussion at the end of the workshop. The across-grade discussion is described later in this report.

Workshop Staff

DRC, QAI, and ACS organized the performance level setting on behalf of DESE. The vendors collaborated to provide general facilitation, including participant training.

DRC facilitated the performance level setting for grades 5 and 8 science. Ricardo Mercado, DRC Research Director, provided general facilitation. Christie Plackner, DRC Director of Research Quality and Data Forensics, facilitated the workshop for grade 5 science; and Joanna Tomkowicz, DRC Research Scientist, facilitated the workshop for grade 8 science. Dave Durette, DRC Sr. Director of Test Development, provided content expertise to participants in grades 5 and 8 science. Sara Kendallen, DRC Sr. Research Analyst, led the analysis team for grades 5 and 8; Ms. Kendallen was supported by Ping Wan, DRC Psychometrics Director, and Julie Pointner Korts of DRC Psychometric Services. Lindy Wienand, DRC Sr. Director of State Programs, coordinated the workshop logistics. Before the workshop, Mr. Mercado, Ms. Kendallen, and Ms. Pointner Korts prepared the physical workshop materials. After the workshop, this team contributed to this report.

ACS facilitated the performance level setting for Physical Science and Biology on behalf of QAI. Andrew Wiley, ACS Partner, provided general facilitation and also facilitated the performance level setting for Physical Science. Deborah Schnipke, ACS Senior Psychometrician, facilitated the performance level setting for Biology. Michelle Udvare, QAI Senior Assessment Specialist, and Les Sewall, QAI Senior Assessment Specialist, provided content expertise to participants in Physical Science and Biology. Dr. Wiley and Dr. Schnipke were supported by Adam Johnson, QAI Program Manager, and Wonsuk Kim, QAI Senior Psychometrician, and Håkan Bergon, QAI State Director, who provided technical and logistical coordination for the performance level setting. Before the workshop, Dr. Wiley, Dr. Kim, and Mr. Johnson prepared the physical workshop materials. After the workshop, Dr. Wiley and Dr. Kim contributed to this report.

DESE Representation

Staff members from DESE oversaw the workshop, provided policy-based guidance to participants when asked, and began the workshop with a brief opening session. DESE's representatives were led by Lisa Sireno, DESE Standards and Assessment Administrator, working in collaboration with Shaun Bates, DESE Director of Assessment, and Kristen McKinney, DESE Director of Science.

DESE was also represented at the workshop by Debbie Jameson, DESE Director of English Language Arts, Lisa Scroggs, DESE Assistant Director of English Language Arts, Chip Sharp, DESE Director of Mathematics, and Dixie Grupe, DESE Director of Social Studies. DESE's representatives monitored the workshop, including the conversations of participants, and responded to questions when called upon.

DESE was assisted in oversight by a member from the technical advisory committee (TAC), Ron Mertz. Dr. Mertz observed the workshop and provided feedback to DESE after the workshop on the proceedings.

Workshop Materials

Participants studied five key pieces of information at the workshop: the Missouri Learning Standards (MLS), performance level descriptors (PLDs), ordered item booklets (OIBs), item maps, and benchmarks.

Missouri Learning Standards (MLS)

Participants were given copies of the MLS for their assigned grade and content area combination. The MLS show what students should be taught and learn in each grade. Participants were instructed that their cut score recommendations should ultimately be tied to the content described in the MLS.

Ordered Item Booklets (OIBs)

A separate OIB was prepared for each test. Item difficulty was calculated using data from Missouri students' performance on the tests. Easier items appeared earlier in the OIB, and harder items appeared later. Items ascended in terms of difficulty throughout the OIB. Multiple choice (MC) and constructed-response (CR) items were ordered together in the OIB. For CR items, each non-zero score point was ordered separately. For example, a two-point CR item appeared twice in the OIB: once for the first score point, and once for the second.

Table 1 shows the total number of pages in each OIB used at the performance level setting.

For grades 5 and 8 science, each OIB comprised items selected from both operational cores of the test. Each OIB was augmented with 3–5 additional items, as taken from the 2018 science field test, to promote coverage along the test scale (i.e., to have a collection of easy, medium, and difficult items). The items in each OIB were selected so they matched the test blueprints: the content coverage of the items in the OIB mirrored that of the student test.

For Biology, the OIB comprised items selected from both operational test forms. All operational items from the two forms were calibrated and used to develop the OIB. For Physical Science, only 1 test form was used in the spring of 2018. As a result, in order to build the OIB with a sufficient number of items to cover the expected range of difficulty, items that were pretested in the spring of 2018 were also used to create the OIB. For the Physical Science test, approximately 37% of the items in the OIB were pretest items.

Table 1. Number of pages in each ordered item booklet, by test

Test	Number of Pages in the Ordered Item Booklet (OIB)
Grade 5 Science	70
Grade 8 Science	70
Physical Science	81
Biology	86

Response probability (RP) criterion. To order the items, an RP criterion was applied. At the 2016 and 2018 MAP performance level settings and at the SBAC performance level setting, an RP criterion of 0.50 was used, often abbreviated as RP50. When RP50 is applied, the RP-adjusted scale location for an item is defined as the scale value associated with a 50% chance of answering the item correctly. For a score point, it is the value associated with a 50% chance of scoring at least that many points.

Selecting an RP criterion represents a policy decision. Various other states and agencies have selected RP criteria such as RP67 with an adjustment for guessing (see Cizek & Bunch, 2007). However, DESE noted that the selection of RP for the 2019 MAP performance level settings should be consistent with that used for ELA and mathematics. Accordingly, DESE selected an RP criterion of RP50. This RP criterion was applied to all four OIBs and item maps.

Item Maps

The item map presented information for the items in the OIB. The item map showed each item's rank-order difficulty, RP-adjusted scale location (difficulty), scoring key, and the aligned content standard.

Benchmarks

Benchmarks were presented for participants' consideration as they recommended their *Proficient* cut scores. The calculation of the benchmarks (and *benchmark ranges*) is described earlier in this section.

As previously stated, DESE indicated that the *Proficient* cut scores on MAP should be informed by the benchmarks (e.g., NAEP for grades 5 and 8). However, when presented to participants, DESE and facilitators made it clear that the benchmarks were provided for participants' consideration only: participants were free to recommend any cut scores that were consistent with the content-based expectations for students, even if those cut scores were outside the benchmark ranges.

Participants were told that if they recommended cut scores in the *Proficient* range, the percentage of students classified at or above *Proficient* on MAP would be similar to the benchmarks. Table 2 shows the *Proficient* benchmark (and, in parentheses, the *Proficient* range) associated with each OIB. For EOC, a single Proficient benchmark was not provided. The OIB range provided to panelists is provided in the table without a single benchmark identified.

Table 2. Proficient benchmark (and OIB pages designated as the Proficient *range), by test*

Test	Benchmark (and Proficient Range)
Grade 5 Science	48 (39 – 54)
Grade 8 Science	46 (40 – 53)
Physical Science	(50 – 69)
Biology	(51 – 75)

For grades 5 and 8 science, participants were given copies of the NAEP *Proficient* achievement level descriptors (ALDs) from grade 4 or grade 8, respectively. During the workshop, facilitators drew participants' attention to the fact that there were important differences in the content and focus between NAEP *Proficient* and MAP *Proficient*, and the facilitators led participants in a discussion of the most salient similarities and differences.

Workshop Procedure

All participants convened in a single room for the opening session, led by Ms. Sireno and Mr. Bates of DESE. DESE set the tone for the workshop by describing the development of the MAP assessments, noting DESE's expectations for appropriately rigorous cut score recommendations. To this end, DESE informed participants that they would be shown benchmarks for their consideration, and that they would work with their colleagues to discuss the knowledge, skills, and abilities that were expected of students in each performance level. DESE also summarized the process that would be used to review and approve the cut points. DESE informed participants that they would work individually and in concert to make recommendations, and these recommendations will be considered by DESE and by other groups.

The general facilitators then provided the first of two whole-committee training sessions. Dr. Wiley and Mr. Mercado summarized again the purpose of the workshop and described the performance level setting method. Participants were given training versions of the PLDs, OIB, and item map to review. DRC described the roles of facilitators, table leaders, and participants to the committee. Participants were told how they would study the OIBs, consider the benchmarks, discuss the content-based expectations for students in each performance level, and make cut score recommendations using the Bookmark Procedure. Participants were reminded to keep the workshop materials confidential. Participants then moved to their pre-assigned groups and tables.

Discussing the Threshold Students. Within each breakout room, the facilitator welcomed participants and led informal introductions. To begin the performance level setting process, the group discussed the MLS, PLDs, and *threshold students*. The threshold students are the hypothetical students with ability at the point-of-entry of each performance level. Participants saw how the knowledge, skills, and abilities detailed in the MLS were reflected in the PLDs. In their tables, participants discussed their content-based expectations for the threshold students. Participants then discussed the threshold students across tables. Each group informally recorded their expectations of the threshold students in bulleted lists which were later posted on the walls of the breakout rooms.

Reviewing the Online Test Participants then reviewed the online test. Participants were given the opportunity to take an operational form of the test in a non-operational, simulated (but realistic) testing environment. By experiencing the online test, participants better understood how the test was administered, how students interacted with the technology-enhanced test items, and how students navigated the test.

Introduction to the Proficient range. DRC then introduced participants to the benchmarked *Proficient ranges*. Participants were told how the benchmarks were calculated and what they represent.

For grades 5 and 8 science, participants reviewed copies of the *Proficient* ALDs from NAEP, and participants discussed the similarities and differences between the MAP PLDs and NAEP ALDs. Participants were reminded that DESE anticipated that their *Proficient* bookmarks may be in the *Proficient* range, so they should pay special attention to this range when studying the OIB.

For Physical Science and Biology, participants were reminded that the benchmark committee did not feel strongly that the *Proficient* bookmark would be within the *Proficient* range, but that this range gave participants a good starting point for their content-based exploration of the OIB.

Studying the OIB. At their tables, participants then studied the items in the OIB. On their item maps, participants noted the content measured by each item or score point. By studying the items in the OIB, participants built an understanding of the knowledge and skills measured by the test.

Refresher Training. Participants were then convened in a refresher training session. During this training, Mr. Mercado and Dr. Wiley reminded participants how bookmarks could be placed in the OIB to represent cut scores. Participants were instructed to keep the threshold students and *Proficient* ranges in mind as they placed their bookmarks.

For grades 5 and 8 science, participants were reminded that the *Proficient* ranges were provided for their consideration, and that the center of the *Proficient* range was associated with the NAEP benchmark. DRC noted that if participants placed their bookmarks in the center of the range, the percentage of students classified as *Proficient* or above on MAP would likely be comparable with the percentage of Missouri students classified as *Proficient* or above on NAEP. However, participants were reminded that their primary task was to make bookmark placements in the OIB that were consistent with the PLDs, with the tested content, and with their expectations for students. Participants were instructed to consider the *Proficient* range as they placed their *Proficient* bookmarks, and then to make bookmark placements for *Advanced* and for *Basic*.

For Physical Science and Biology, panelists were reminded that the *Proficient* range was provided to offer some additional context for the panelists' rating. They were also reminded about the concerns raised by the Benchmark panel and that there was not an explicit expectation that their ratings would be consistent with the *Proficient* benchmark range. As with the grade 5 and grade 8 panels, participants were instructed to consider the *Proficient* range as they placed their *Proficient* bookmarks, and then to make bookmark placements for *Advanced* and for *Basic*.

To gain insight into participants' rationales behind their bookmarks, participants were asked to write down their *content-based rationales* on the judgment form they used to record their bookmark placements. Participants were shown sample rationales that took the content measured by items in the OIB around the bookmarks; and linked this content to the content-based expectations for the threshold students.

Participants were then given an evaluation to gauge their understanding of the performance level setting process, its materials, and the process to that point. After answering participants' questions on bookmark placement, participants returned to their tables to begin the Round 1 of the Bookmark Procedure.

Round 1. Individually, participants considered the knowledge, skills, and abilities measured by the items in the OIB. Participants were reminded that their primary task was to place bookmarks that are consistent with the MLS, the PLDs, the tested content, and the expectations for the threshold students.

In addition, for their *Proficient* bookmarks, participants were instructed to consider the *Proficient* range. To do so, participants were instructed to use the process shown here.

1) Consider the content-based expectations of the threshold *Proficient* student.
2) Start with a bookmark placement in the middle of the *Proficient* range, and consider the knowledge and skills measured by the items contained before that bookmark placement.
3) Determine whether there is good correspondence between the content measured by the items before their bookmark and the expectations for the threshold *Proficient* student.
4) If there is good correspondence, keep that bookmark placement. If there is not good correspondence, move the bookmark forward or backward in the OIB (one page at a time) until there is good correspondence.

Participants recorded content-based rationales for each of their decisions, explicitly linking the content measured by the items before the *Proficient* bookmark with their conceptualization of the threshold students. Participants were reminded that their *Proficient* bookmarks may fall within the *Proficient* range, but that this was not compulsory.

After placing their *Proficient* bookmark, participants placed their *Advanced* and *Basic* bookmarks and recorded content-based rationales. Participants made their Round 1 judgments without discussion.

Workshop staff tabulated participants' Round 1 bookmark placements and calculated each group's median cut score recommendations. The group's median cut score recommendation was taken as the recommendation of the committee. The Round 1 results for each group are presented in Section F of this report.

Facilitators presented participants with feedback based on their Round 1 bookmark placements, including (a) a histogram showing the bookmarks placed in Round 1, and (b) the group's median bookmark placements. Facilitators discussed the variability of participants' Round 1 *Proficient* bookmarks and how they compared with the *Proficient* range. Facilitators reminded participants about the recommended interpretation of the *Proficient* range and how the group should consider it when placing *Proficient* bookmarks. Participants were instructed that it was normal and expected to have significant variability between participants when considering their Round 1 bookmarks, and that they would have an opportunity to discuss their bookmarks with their colleagues in Rounds 2 and 3.

In their tables, participants then shared their Round 1 bookmark placements and content-based rationales with their colleagues. Participants were encouraged to refer to the OIB, item map, MLS, PLDs, benchmarks, and threshold student expectations throughout this discussion.

Round 2. Following the discussion, participants again individually considered their bookmark placements. All participants made their bookmark placements individually and without discussion.

Workshop staff tabulated participants' Round 2 bookmark placements, calculated each group's median cut score recommendations, and calculated *impact data*. Impact data are the percentages of students who would be classified in each performance level given a set of cut scores.

Facilitators presented participants with feedback based on their Round 2 bookmark placements, including (a) a histogram showing the bookmarks placed in Round 2, (b) the group's median bookmarks, and (c) the impact data. The Round 2 results for each group are presented in Section F of this report.

Facilitators then led participants in a discussion of their Round 2 recommendations across both tables. Participants were asked to share the rationales behind their bookmark recommendations. Participants were asked to refer frequently to the workshop materials, such as the OIB and PLDs. Participants were reminded that they should keep using the MLS and PLDs when considering the knowledge, skills, and abilities that each of the three threshold students should be able to show command of; and that there may sometimes be a difference between what a student *should* be able to do, versus what he or she can *currently* do given his or her level of preparation and instruction.

Round 3. Participants then made their Round 3 bookmark placements. All participants made their bookmark placements individually and without discussion. Once complete, the facilitator presented the group with feedback based on Round 3 (i.e., median bookmarks, histogram of bookmarks, impact data based on Round 3 median cut score recommendations).

Participants from each group then reconvened in a general session. The general facilitators presented the committee with the impact data associated with the cut score recommendations for all four tests. Participants were asked to consider how reasonable they found the pattern of impact data across grades. Participants were asked to discuss (a) whether they would recommend adjustments to their group's cut scores to promote better *articulation* across grades; (b) if needed, how far the group could recommend adjusting its median bookmark placements while still being consistent with the MLS, PLDs, and threshold student expectations; and (c) if the group had questions for other groups about their cut score recommendations. Table 3 shows the recommended cut scores from Round 3 of the performance level setting, plus the associated impact data using spring 2019 administration data.

Table 3. Cut scores and associated impact data from Round 3 of the grade-level performance level setting

	Round 3 Cut Scores			**% Students by Level Based on Spring 2019**				
Test	***Basic***	***Proficient***	***Advanced***	***Below Basic***	***Basic***	***Proficient***	***Advanced***	***Prof. + Adv.***
5	275	310	344	26.0%	31.1%	29.3%	13.5%	*42.8%*
8	461	501	537	16.5%	30.4%	35.0%	18.1%	*53.1%*
PHY	382	400	417	19.2%	42.9%	28.2%	9.7%	*37.9%*
BIO	381	400	406	16.8%	43.8%	14.1%	25.4%	*39.5%*

Participants in each group considered the pattern of cut scores and impact data across grades. Participants were instructed to tell their table leaders about their views on the across-grade articulation of the cut scores; and in turn, table leaders were asked to make notes about their participants' opinions.

Participants then evaluated the workshop: the results of the evaluation are discussed later in this section under the heading "Workshop Evaluation." Table leaders then convened in the across-grade discussion, and the remaining participants were dismissed with the thanks of DESE and the facilitators.

Across-grade discussion. Table leaders from each group were convened to examine the articulation of the recommendations across grades and tests.

The group was given time to consider the consistency of the recommendations across grades and tests. In this sense, performance standards are consistent when the impact data form a meaningful, explainable pattern across tests. In particular, each group was asked to consider the extent to which each grade's recommendations were consistent with the others across the four tests of science.

After a brief introduction, the facilitators told the group that only table leaders from the group who worked on a test could recommend adjustments to those cut scores (e.g., only grade 8 science table leaders could recommend adjustments to the grade 8 science cut scores). DRC also stressed that table leaders should only recommend cut scores that are consistent with the tested content.

Table leaders shared their participants' thoughts on the pattern of the impact data across grades. The table leaders reiterated that the small population taking the Physical Science test made it difficult to compare those impact data directly with the other tests. For the remaining tests (i.e., grades 5 & 8 science and Biology), the table leaders noted that they expected the percentage of students classified in each performance level to be fairly consistent across grades. However, the table leaders noted for grade 5 science, where students may receive little science instruction prior to grade 5, the percentage of students classified as *Proficient* and above might be expected to be somewhat lower than other grades.

The table leaders for grade 8 science noted that the percentage of students classified as *Below Basic* was lower than observed in other grades. The other table leaders concurred with this assessment. Working together, the table leaders for grade 8 science suggested that the *Basic* cut score for grade 8 science be raised from a bookmark of 15 (associated with a cut score of 461) to a bookmark of 17 (associated with a cut score of 468, a difference of 7 scale score points or +0.6 SEM). This bookmark was still consistent with the group's content-based recommendations, and the articulation of the percent of students classified as *Basic* and above across grades was improved.

The table leaders for Biology noted that the percentage of students classified as *Advanced* was unexpectedly high for their test. They reported that their group had not expected this percentage to be so high, and that small changes to their median bookmarks had changed the percentage of students classified as *Advanced* a great deal. Working together, the table leaders for Biology recommended raising the *Advanced* cut score for Biology from a bookmark of 83 (associated with a cut score of 406) to a bookmark of 84 (associated with a cut score of 411, a difference of 5 scale score points. This bookmark was consistent with the group's content-based expectations for students in the *Advanced* performance level, and the articulation of the impact data across grades was improved.

The facilitators asked whether the table leaders wished to consider any additional adjustments to the recommendations to promote across-grade articulation. The group noted that the percentage of students classified as *Proficient* and above was higher in grade 8 than any other grade, and that this pattern was not necessarily expected. However, the group did not come to a consensus to make a recommendation to change this cut score.

Working by consensus, the group recommended two changes to cut scores, as detailed above: the grade 8 *Basic* cut score and to the Biology *Advanced* cut score. Both changes were suggested to promote consistency across grades.

Table 4 shows the recommended cut scores from the performance level setting across-grade discussion, plus the associated impact data using spring 2019 administration data.

Table 4. Recommended cut scores and associated impact data from the performance level setting across-grade discussion

	Articulated Cut Scores			% Students by Level Based on Spring 2019				
Test	***Basic***	***Proficient***	***Advanced***	***Below Basic***	***Basic***	***Proficient***	***Advanced***	***Prof. + Adv.***
5	275	310	344	26.0%	31.1%	29.3%	13.5%	*42.8%*
8	468	501	537	20.4%	26.5%	35.0%	18.1%	*53.1%*
PHY	382	400	417	19.2%	42.9%	28.2%	9.7%	*37.9%*
BIO	381	400	411	16.8%	43.8%	24.5%	15.0%	*39.5%*

Standard Errors Associated with Recommendations

Statistical uncertainty (or noise) accompanies any cut score. One related statistic is shown here.

1) **Standard error of measurement (SEM)** quantifies the measurement error associated with the test instrument at the point on the test scale equal to the cut score.

These values, associated with each cut score, are shown in Table 5. In the table, the SEM refers to the conditional SEM associated with the cut scores presented in Table 4. All values are in scale-score units.

Table 5. Standard error values associated with the cut score recommendations, by test

	Standard Error of Measurement (SEM)		
Test	***Basic***	***Proficient***	***Advanced***
5	11	11	12
8	11	9	9
PHY	5	5	6
BIO	4	4	5

Workshop Evaluation

Participants completed a written evaluation of the performance level setting. In addition, table leaders completed a written evaluation of the across-grade discussions. Results from the evaluations can be used to gauge how fair and valid the participants perceived the performance level setting process to be, and whether participants supported their cut score recommendations.

Of the 47 participants, 45 completed evaluations at the end of the workshop. The evaluation results showed that participants understood the process and supported their recommendations. Selected results are shown here.

- 100% understood the purpose of the workshop.
- 98% understood what the benchmarks were and how they should be considered.
- 100% believed their opinions were considered and valued by their group.
- 87% saw their group's work reflected in the presentation of recommendations across grades.
- 91% believed this process will lead to defensible performance standards for the test.
- 93% were comfortable with the group's recommendation for the *Basic* cut score.
- 78% were comfortable with the group's recommendation for the *Proficient* cut score
- 82% were comfortable with the group's recommendation for the *Advanced* cut score

Of the eight table leaders, all eight completed evaluations at the end of the across-grade discussion. The evaluation results showed that table leaders respected their groups' recommendations and supported their ultimate recommendations. Selected results are shown here.

- 100% understood the purpose of the across-grade articulation.
- 100% considered the recommendation from their original group during the discussion.
- 100% considered the content-based expectations for students during the discussion.
- 88% believed this process will lead to defensible performance standards for the test.
- 100% felt the *Basic* cut score for their own grade was about right (i.e., not too high or low).
- 88% felt the *Proficient* cut score for their own grade was about right.
- 88% felt the *Advanced* cut score for their own grade was about right.
- 100% felt the *Basic* cut score for the other grades in the content area was about right.
- 63% felt the *Proficient* cut score for the other grades in the content area was about right.
- 100% felt the *Advanced* cut score for the other grades in the content area was about right.

Taken as a whole, the results of the workshop evaluations show that participants were generally satisfied with the performance level setting and with their recommendations. However, the results of the evaluation from the across-grade discussion indicated a minority of table leaders had lingering questions about the *Proficient* cut scores, particularly for grade 8 science. These concerns were addressed by DESE after the performance level setting.

After the Performance Level Setting Workshop

After the performance level setting, DESE reviewed the cut score recommendations from Missouri educators. The Department acknowledged the amount of time and attention spent by Missouri educators in making well-reasoned, content-based cut score recommendations for the assessments.

DESE noted that the percentages of students classified as *Proficient* and above in grade 8 science was higher than that observed on other tests. At the across-grade discussion, several table leaders pointed out this pattern; however, the committee of table leaders could not make a consensus recommendation to adjust the *Proficient* cut score for grade 8 science.

In previous consultations, DESE's technical advisory committee (TAC) has noted that performance level setting committees use content-based information to make their recommendations, and that DESE had the responsibility to recommend cut scores which were reasonable from both a content- and a policy-based standpoint. The TAC has indicated that a very common and defensible practice by state departments of education is to take participants' recommendations and make post-workshop adjustments in order to promote better articulation across grades and tests. By taking such an approach, DESE can:

a) honor the voice of Missouri educators who participated at the performance level settings,
b) use Missouri students' test data to make adjustments that promoted good articulation across grades and tests, and
c) continue to adhere to industry best practices.

The TAC had previously advised that one approach to making such adjustments uses standard error of measurement (SEM). When a student takes any test, a certain amount of statistical error is associated with his or her test score. This statistical error can be quantified by SEM. If a student were to be retested (without remembering the previous test's questions), one would expect that his or her second score would be similar to—but not exactly equal to—the first score. Were the student retested many times, one would expect that his or her scores would form a distribution, and that the scores would fall within a band of ±1 SEM about two-thirds of the time. As such, differences in test scores of less than ±1 SEM are not considered meaningful, as they fall within a range of statistical error quantified by this statistic.

DESE considered educators' recommendations, and it considered the policy-based ramifications of implementing the cut scores. DESE noted that having well-articulated performance standards across all grades and tests was important to the MAP testing program. At the same time, DESE noted that performance level setting participants had used the content-based information available to them to recommend well-reasoned cut scores during the performance level setting. Accordingly, DESE made the decision to consider adjusting participants' recommended cut scores using SEM. The adjustment shown in Table 6 was considered.

Table 6. Post-workshop adjustments to cut scores, by test

Test	Cut Score	Adjustment
Grade 8 Science	*Proficient*	+1 SEM

Although it reserved the right to do so, DESE proposed no adjustment to the other 11 cut scores recommended by Missouri educators: the other cut scores were left unadjusted. Table 7 presents the MAP science cut scores and associated impact data from DESE's review of the recommended cut scores.

Table 7. Cut scores and associated impact data from DESE's review of the recommended cut scores for MAP science

	Recommended Cut Scores			% Students by Level Based on Spring 2019				
Test	***Basic***	***Proficient***	***Advanced***	***Below Basic***	***Basic***	***Proficient***	***Advanced***	***Prof. + Adv.***
5	275	310	344	26.0%	31.1%	29.3%	13.5%	*42.8%*
8	468	510	537	20.4%	35.7%	25.9%	18.1%	*43.9%*
PHY	382	400	417	19.2%	42.9%	28.2%	9.7%	*37.9%*
BIO	381	400	411	16.8%	43.8%	24.5%	15.0%	*39.5%*

Approval of the Cut Scores

To approve the cut scores, DESE has indicated it will complete three important steps:

- **Consultation with DESE leadership.** DESE will consider the cut scores and associated impact data, and it will continue to consider the potential impact of the implementation of the science cut scores on Missouri students and on the MAP testing program.
- **Consultation with the TAC.** DESE will consult with its technical advisory committee (TAC) in August 2019. During this meeting of the TAC, DESE and DRC, QAI, and ACS will present the results of the performance level setting to this committee of statewide- and nationally-recognized experts in educational measurement.
- **Presentation to the Missouri State Board of Education.** In Fall 2019, DESE will present the final cut score recommendations to the Missouri State Board of Education. The Board will be invited to act on the cut scores as they are implemented for grade 5 science, grade 8 science, Physical Science, and Biology.

References

Buckendahl, C.W. & Wiley, A. (2016). *Performance level descriptor development workshop for the Missouri Department of Elementary and Secondary Education: Summary report*. Las Vegas, NV: ACS Ventures.

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests.* Thousand Oaks, CA: Sage.

Data Recognition Corporation. (2016). *Missouri Assessment Program grades 3–8 English language arts and mathematics: Standard setting 2016 final technical report.* Jefferson City, MO: DESE.

Data Recognition Corporation. (2018). *Missouri Assessment Program grades 3–8 English language arts and mathematics: Standard setting 2018 final technical report.* Jefferson City, MO: DESE.

Egan, K.L., Schneider, M.C., & Ferrara, S. (2012). Performance level descriptors: History, practice, and a proposed framework. In G. J. Cizek (Ed.*), Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 79-106). New York: Routledge.

Karantonis, A., & Sireci, S. G. (2006). The Bookmark standard setting method: A literature review. *Educational Measurement: Issues and Practice, 25*, 4–12.

Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996). *Standard setting: A Bookmark approach*. Symposium presented at the Council of Chief State School Officers National Conference on Large-Scale Assessment: Phoenix, AZ.

Lewis, D. M., Mitzel, H. C., Mercado, R. L., & Schulz, E. M. (2012). The bookmark standard setting procedure. In G. J. Cizek (Ed.*), Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 225-253). New York: Routledge.

McClarty, K.L., Way, W.D., Porter, A.C., Beimers, J.N., & Miles, J.A. (2013). Evidence-based standard setting: Establishing a validity framework for cut scores. *Educational Researcher (42)*, 2, 78–88. https://doi.org/10.3102/0013189X12470855

Missouri State Board of Education. (2016, April 19). *Minutes of meeting of the State Board of Education: April 19, 2016.* Retrieved from https://dese.mo.gov/sites/default/files/April2016Minutes.pdf

Missouri Department of Elementary and Secondary Education. (2017). *Grade-Level: Achievement level descriptors*. Retrieved from https://dese.mo.gov/college-career-readiness/assessment/grade-level.

Phillips, G.W. (2012). The benchmark method of standard setting. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 323-346). New York: Routledge.

C
Agenda

Workshop Agenda

Department of Elementary and Secondary Education (DESE)

Missouri Assessment Program (MAP)
Grades 5 & 8 Science, Physical Science, and Biology

Standard Setting Workshop
Columbia, MO
July 16–18, 2019

Welcome to the standard setting workshop for the Missouri Assessment Program (MAP) tests of Grades 5 and 8 Science, Physical Science, and Biology! The Missouri Department of Elementary and Secondary Education (DESE), Data Recognition Corporation (DRC), Questar, and ACS would like to thank you for your time and expertise during this important process.

Please use this agenda to orient yourself during the workshop. If you have any questions or concerns, please do not hesitate to contact a member of the workshop staff.

Tuesday, July 16

Welcome!

7:30 AM	**Participant Registration and Breakfast** Participants check in at the reception table to sign in, receive a nametag, and collect any other necessary information.
8:30 AM	**Opening Session** DESE welcomes participants, overviews the MAP assessments, discusses the reasons for the standard setting workshop, and describes the desired outcomes of the workshop.
9:00 AM	**Participant Training** The facilitators introduce participants to the Bookmark Standard Setting Procedure, explaining how a cut score can be represented in an ordered item booklet (OIB) as a bookmark.
10:15 AM	**Break**
10:30 AM	**Distribution of Secure Materials in Breakout Rooms** Participants reconvene in their pre-assigned groups and breakout rooms. After brief introductions, facilitators distribute the secure workshop materials. • Please remember that the secure materials must remain in your breakout room and that your discussions of the secure materials must remain confidential. • Be sure to write your name on each of the secure materials. • Your folder number is printed on the upper-right of your workshop folder. Write your name and folder number on the materials sign-out list.

Tuesday, July 16 *(continued)*

Threshold Students Discussion and Student Test

10:45 AM **Discuss the Standards, PLDs and Threshold Students**

As part of a structured discussion, participants discuss the Missouri Learning Standards (MLS), performance level descriptors (PLDs), and threshold students.

- On your own, take 10–15 minutes to review the Standards and PLDs for your grade and content area. Feel free to write on your copies of these materials.
- Read the PLDs to examine the knowledge, skills, and abilities that *Proficient* students are expected to have. Then do the same for *Advanced* and *Basic*.
- As a table, take 30 minutes to discuss the content-based expectations for the threshold students (i.e., the students just barely entering each performance level). Remember that there are three threshold students to consider: just *Basic,* just *Proficient* and just *Advanced*.
- The table leader should appoint a scribe to take notes during the discussion. For each threshold student, the table should help the scribe create a brief, bulleted list on easel paper that describes the skills expected of that threshold student.
- Using the remaining time, the facilitator will ask the table leaders to "report out" on the discussion, sharing with the other table the knowledge, skills, and abilities expected of each threshold student. Scribes should update their table's bulleted lists of the content-based expectations of each threshold student.

Noon **Lunch**

The group breaks for 45 minutes.

12:45 PM **Review the Student Test**

Participants review the student test to get a sense of what students saw on test day.

- On your own, take the test at a computer station.
- When taking the test, briefly examine the items to get a general sense of what is measured by the test and how it is measured.
- Briefly examine the item types and the test tools to get a general sense of item functionality and how the tools are used to aid the student in answering items.
- If necessary, use the provided index cards to record comments about test items.

Tuesday, July 16 *(continued)*

Ordered Item Booklets (OIBs)

1:45 PM **Break**

2:00 PM **Introduction to the Ordered Item Booklet (OIB) and Presentation of Benchmarks**
Facilitators introduce this task by instructing participants to find the OIB and item map in their secure materials. Facilitators then present the benchmarks for participants' consideration.

- Benchmarks are based on the performance of Missouri students on the tests.
- Mark the position of the benchmarks in your OIB so you can consider them throughout the standard setting workshop.
- If necessary, use the provided index cards to record comments about test items.

2:15 PM **Begin Study of the OIB**
Table leaders lead their tables in an examination of the items in the OIB.

- Participants should study the first 20 items in the OIB on their own, recording on their item what each item measures.
- Then at their tables, participants engage in a brief discussion about these first items in the OIB. The informal discussion should be led by the table leader.
- Following this brief discussion, the group should continue with the next group of 20 items, discussing them as a table after several minutes of independent study.
- The group should give special attention to the items around the benchmarks.
- Each participant records these details on his or her own copy of the item map.
- Facilitators monitor the tables to check that each participant has a chance to speak.

4:25 PM **Secure Materials Collection**
Facilitators lead the collection of the secure materials from all participants.

- Table leaders help the participants at their table pack up their secure materials in the manner directed by the facilitator.

4:30 PM **Dismissal**
Participants are reminded that the workshop reconvenes at 8:00 am on Wednesday.

Wednesday, July 17

Bookmark Training and Round 1

7:30 AM	**Participant Sign-In and Breakfast** Participants sign-in for the day in their breakout rooms.
8:00 AM	**Complete Study of the OIB** In their tables, participants complete their study of the items in the OIB, paying special attention to the items around the benchmark.
9:30 AM	**Bookmark Placement Training** In a general session, the facilitators reintroduce bookmark placement, explaining and illustrating how bookmarks are placed and what bookmarks mean. • The facilitator explains how participants make cut score recommendations by placing bookmarks in the OIB. • After the training, a brief mid-process evaluation is administered.
10:15 AM	**Break**
10:30 AM	**Round 1 Bookmark Placement** In their breakout rooms, facilitators direct all participants to place their Round 1 bookmarks. • Remember that bookmark placement is an individual activity. • Be sure to consider the benchmarks when placing your *Proficient* bookmark. • Record your bookmark placements and content-based rationale on your bookmark worksheet. Then submit your bookmark worksheet to your facilitator.
11:15 AM	**Presentation of Round 1 Results** The facilitator presents a summary of the bookmark placements made in Round 1.
11:30 AM	**Begin Discussion of Round 1 as a Table** Table leaders facilitate a discussion of the Round 1 judgments at their tables. Participants discuss the items between the lowest and highest bookmarks, explaining the rationales behind their judgments. • Each participant shares where he or she placed his or her *Proficient* bookmark in Round 1, along with the content-based rationale behind the recommendation. • After each participant has had an opportunity to speak, the table moves on to the *Advanced* bookmark, then the *Basic* bookmark.
Noon	**Lunch** The group breaks for 45 minutes.

Wednesday, July 17 *(continued)*

Round 2

12:45 PM **Complete Discussion of Round 1 as a Table**

1:15 PM **Round 2 Bookmark Placement**
Facilitators direct all participants to place their Round 2 bookmarks.

- Remember that bookmark placement is an individual activity.
- Record your bookmark placements and content-based rationale on your bookmark worksheet. Then submit your bookmark worksheet to your facilitator.

1:45 PM **Presentation of Round 2 Results**
The facilitator presents a summary of the bookmark placements made in Round 2.

2:00 PM **Begin Discussion of Round 2 as a Group**
Facilitators lead a discussion of the Round 2 judgments across the tables, including the content-based expectations for students in each performance level. Each table reports on the discussions that occurred at the table in Round 2.

- Each table shares the types of discussions it had regarding the *Proficient* bookmark in Round 2. Participants share their bookmark placements and their content-based rationales, referring to the OIB and PLDs whenever possible.
- After each table has had an opportunity to speak, the group moves on to the *Advanced* bookmark, then the *Basic* bookmark.

3:00 PM **Round 3 Bookmark Placement**
Facilitators direct all participants to place their Round 3 bookmarks.

- Remind participants that bookmark placement is an individual activity.

3:30 PM **Break**
During this break, participants should reconvene in the general session room to await the presentation of Round 3 results.

3:45 PM **Presentation of Round 3 Results in General Session**
The facilitators present a summary of the bookmark placements made in Round 3, as well as the impact data associated with the median bookmark placements.

4:00 PM **Discussion of Round 3 Results**
Participants consider the Round 3 recommendations from all the grades and tests. Participants share their views with their table leader about the entire system of performance standards.

- Facilitators encourage participants to look at the articulation of the performance standards across grades.
- Table leaders take notes on their participants' views about the articulation across grades for use in the articulation discussion.

Wednesday, July 17 *(continued)*

Workshop Evaluation and Dismissal

4:15 PM **Workshop Evaluation**
Participants complete an evaluation of the workshop and recommendations.

4:25 PM **Secure Materials Collection**
Facilitators lead the collection of the secure materials from all participants.

- Table leaders help the participants at their table pack up their secure materials in the manner directed by the facilitator.

4:30 PM **Dismissal**
Table leaders are asked to reconvene at 8:00 am on Thursday for the articulation discussion. All other participants are dismissed with the thanks of DESE, DRC, Questar, and ACS.

Thursday, July 18 (for Table Leaders Only)

Articulation Discussion

NOTE: Only table leaders participate in the across-grade articulation discussion
Only table leaders will take part in the articulation discussion. All other participants will be dismissed from the workshop at 4:30 pm on Wednesday, July 17.

7:30 AM **Table Leader Sign-In and Breakfast**
Please be sure to sign in for the day in your breakout room.

8:00 AM **Begin Articulation Discussion**
The facilitators present the same display of Round 3 results that was shown to all participants at the end of the Bookmark Procedure. Working in groups by content area, facilitators ask the table leaders to consider how well the performance standards are articulated across grades, and whether any cut scores should be adjusted to promote better articulation.

- Any table leader may comment on any cut score and ask questions.
- Only the table leader from the group that made a cut score recommendation may suggest changes to it. Then the entire committee of table leaders will vote on the proposed change.
- All cut scores must have clear links to the Missouri Learning Standards, PLDs, and test items.

10:00 AM **Break**

10:15 AM **Complete Articulation Discussion**

11:45 AM **Workshop Evaluation**
Participants complete an evaluation of the workshop and recommendations.

11:55 AM **Secure Materials Collection**
Facilitators lead the collection of the secure materials from all participants.

- Table leaders help the participants at their table pack up their secure materials in the manner directed by the facilitator.

Noon **Dismissal**
All participants are dismissed with the thanks of DESE, DRC, Questar, and ACS.

Missouri Science Standard Setting Workshop

Participant Agenda

Tuesday, July 16

7:30 AM Participant Registration and Breakfast
8:30 AM Opening Session
9:00 AM Participant Training
10:15 AM Break
10:30 AM Distribution of Secure Materials in Breakout Rooms
10:45 AM Discuss the Standards, PLDs and Threshold Students
Noon Lunch
12:45 PM Review the Student Test
1:45 PM Break
2:00 PM Introduction to the Ordered Item Booklet (OIB) and Presentation of Benchmarks
2:15 PM Begin Study of the OIB
4:25 PM Secure Materials Collection
4:30 PM Dismissal

Wednesday, July 17

7:30 AM Participant Sign-In and Breakfast
8:00 AM Complete Study of the OIB
9:30 AM Bookmark Placement Training
10:15 AM Break
10:30 AM Round 1 Bookmark Placement
11:15 AM Presentation of Round 1 Results
11:30 AM Begin Discussion of Round 1 as a Table
Noon Lunch
12:45 PM Complete Discussion of Round 1 as a Table
1:15 PM Round 2 Bookmark Placement
1:45 PM Presentation of Round 2 Results
2:00 PM Begin Discussion of Round 2 as a Group
3:00 PM Round 3 Bookmark Placement
3:30 PM Break
3:45 PM Presentation of Round 3 Results in General Session
4:00 PM Discussion of Round 3 Results
4:15 PM Workshop Evaluation
4:25 PM Secure Materials Collection
4:30 PM Dismissal

Thursday, July 18 (for Table Leaders Only)

7:30 AM Table Leader Sign-In and Breakfast
8:00 AM Begin Articulation Discussion
10:00 AM Break
10:15 AM Complete Articulation Discussion
11:45 AM Workshop Evaluation
11:55 AM Secure Materials Collection
Noon Dismissal

D

Training Presentation and Materials

Missouri
DEPARTMENT OF ELEMENTARY & SECONDARY
EDUCATION

Missouri Assessment Program
Science Performance Level Setting

July 2019

Purpose

- Provide advice and recommendations.
  - Help DESE establish performance expectations in science.
  - Help DESE establish cut scores for science tests.

When is Standard Setting Necessary?

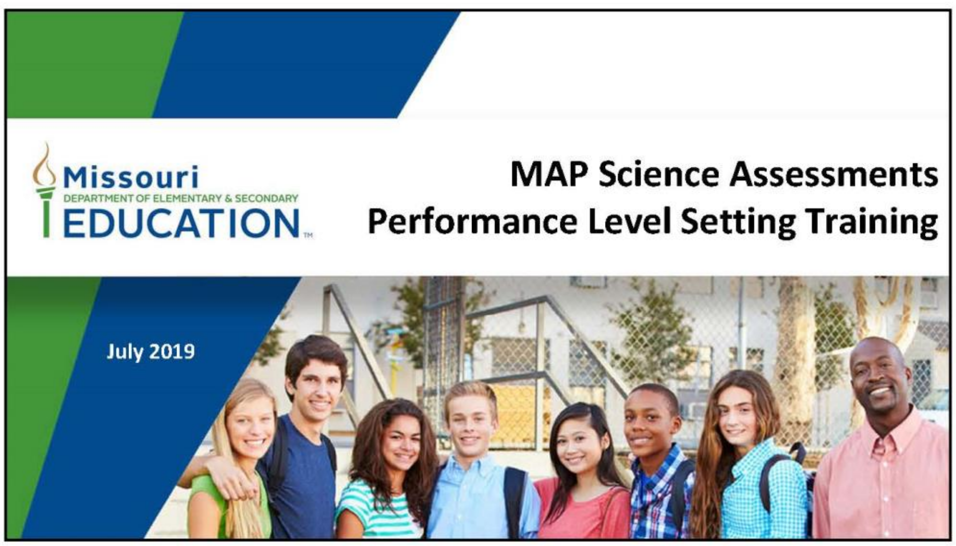
- When academic standards, test design, or content undergo a major change
- When performance expectations change
- When a new test measures something different from the old test
- When previously established cut scores no longer apply

This Week's Work

- Help prepare students for success at the next level.
- Provide information to DESE for the purpose of setting performance levels and corresponding cut scores.
- Consider a variety of information and use it to make content-based decisions.

What Happens Next?

- Statistical and policy review – August
- Official cut scores established – September
- Fall reporting of assessment results – October



Workshop Staff Members

__Data Recognition Corporation__

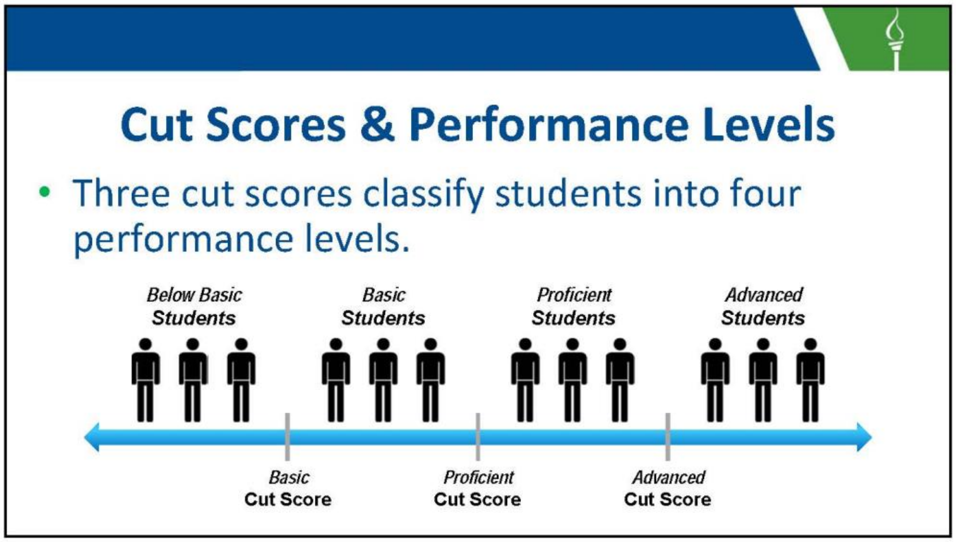
- **Rick Mercado**, General Facilitator
- **Christie Plackner**, Grade 5 Science
- **Joanna Tomkowicz**, Grade 8 Science
- **Dave Durette**, Science Content Expert
- **Sara Kendallen**, Data Analyst
- **Ping Wan**, Data Analyst
- **Lindy Wienand**, Program Manager
- **Mark Sprang**, Information Technology

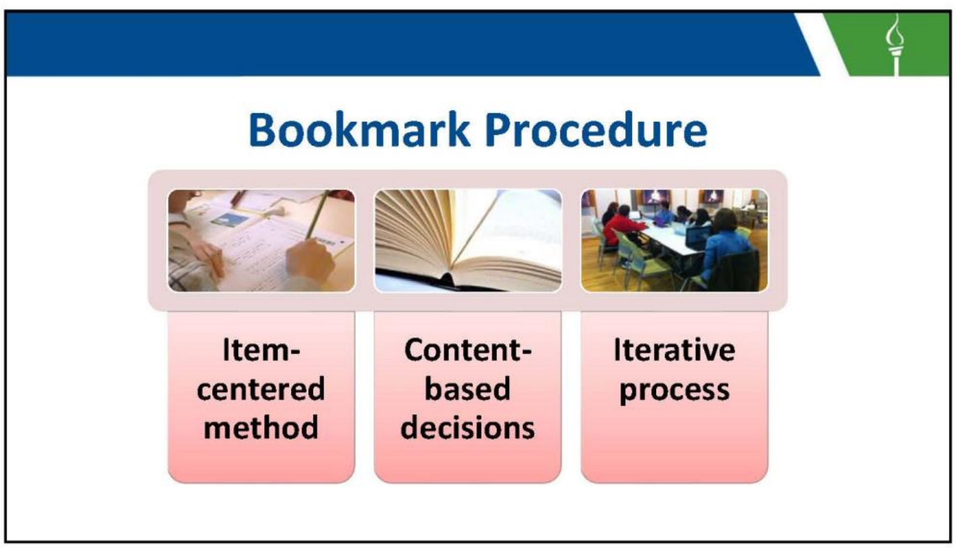
__Questar Assessment & ACS Ventures__

- **Drew Wiley**, Physical Science
- **Deborah Schnipke**, Biology
- **Les Sewall**, Sr. Assessment Specialist
- **Michelle Udvare**, Sr. Assessment Specialist
- **Wonsuk Kim**, Sr. Psychometrician
- **Adam Johnson**, Program Manager
- **Håkan Bergon**, State Director

Workshop Goal

- To ***recommend*** cut scores that categorize students into one of four performance levels:
  - *Below Basic*
  - *Basic*
  - *Proficient*
  - *Advanced*


Cut Scores & Performance Levels
• Three cut scores classify students into four performance levels.
Below Basic Students
Basic Students
Proficient Students
Advanced Students
Basic Cut Score
Proficient Cut Score
Advanced Cut Score


Bookmark Procedure
Item-centered method
Content-based decisions
Iterative process

Process Overview

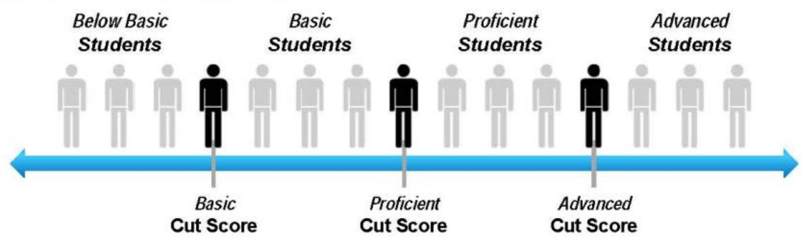
- Study the performance level descriptors
- Discuss the threshold students
- Review the test
- Study the ordered item booklet
- Consider benchmarks
- **Round 1**: Make cut score recommendations on your own
- Discuss your recommendations with your table
- **Round 2**: Make cut score recommendations on your own
- Discuss your recommendations with your group
- **Round 3**: Make cut score recommendations on your own
- Review the group's recommendations
- Evaluate the workshop
- Across-grade discussion for table leaders

Performance Level Descriptors

- PLDs describe the knowledge, skills, and abilities expected of students in each performance level.
  - They are linked to the Missouri Learning Standards.
  - PLDs describe students in the middle of each level, not on the *thresholds*.

PLDs and Performance Levels

- PLDs describe the student in the middle of each performance level.



Three Threshold Students

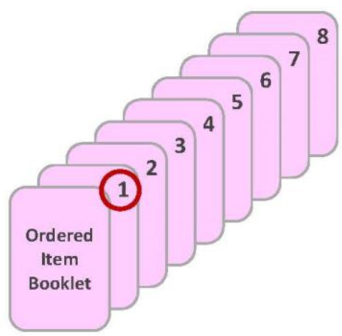
- Threshold students are those just barely leaving one level and entering the next level.
  - The PLDs do *not* describe these students directly.
  - There are three threshold students.

Take the Test

- By taking the test, you will better understand students' testing experience on test day.
- After taking the test on computer, you will see the items on paper, ordered by difficulty, in the ordered item booklet (OIB).

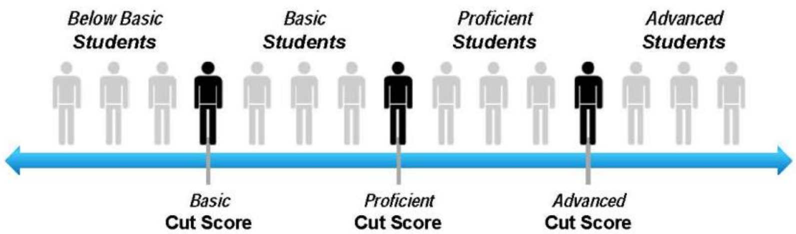
Ordered Item Booklet (OIB)

- The OIB comprises items from the spring test.
  - One item per page
  - Easiest item first
  - Hardest item last
  - Items ascend in difficulty as based on student performance



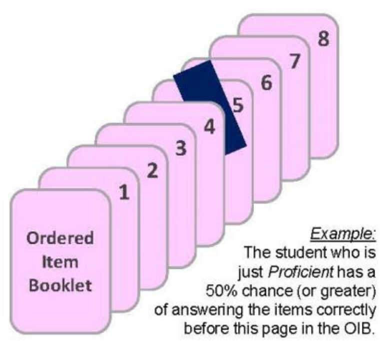
Three Threshold Students

- Bookmark judgments and cut scores are linked to the student *just* in each level.



Threshold Students and the OIB

- You will consider the three threshold students.
- You will make statements in the OIB using *bookmarks.*
- These bookmarks are linked to cut score recommendations.



Three Rounds

Round 1
- Study OIB and make your own bookmark judgments
- Discuss your ratings with your tablemates

Round 2
- On your own, make your own bookmark judgments
- See feedback and discuss your ratings with your group

Round 3
- On your own, make your own bookmark judgments
- Then review recommended cut scores

Roles and Responsibilities

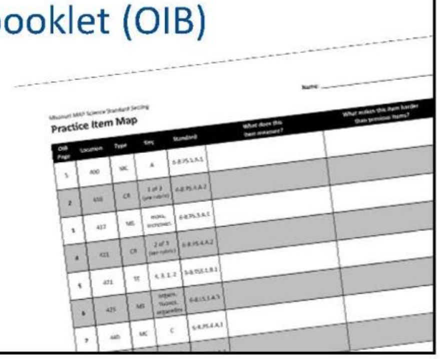
- You will recommend performance standards to DESE.
- During the workshop, remember to:
  - Contribute to discussions at your table
  - Participate in group-wide discussions
  - Place your bookmarks independently
  - Ask a member of staff any questions
  - Use workshop materials only in meeting rooms
  - Keep workshop conversations confidential

Workshop Security

- Your facilitators will collect your materials each afternoon in a structured way.
- Always leave the workshop materials in the meeting rooms. Do not discuss the contents of the materials outside your meeting room.
- You are welcome to use phones, tablets, and laptops in the lunchroom and hallways, but never in the meeting rooms.

Training Materials

- ❑ Training ordered item booklet (OIB)
- ❑ Item map
- ❑ Item separation chart
- ❑ Bookmark worksheet



Item Map

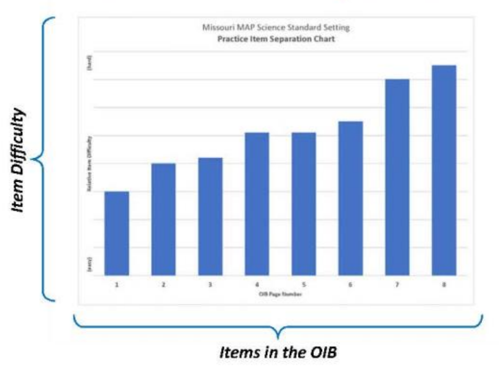
Missouri MAP Science Standard Setting

Practice Item Map

Name: ______________________

OIB Page	Location	Type	Key	Standard	What does this item measure?	What makes this item harder than previous items?
1	400	MC	A	6-8.PS.1.A.1		
2	410	CR	1 of 3 (see rubric)	6-8.PS.4.A.2		
3	412	MS	mass, increases	6-8.PS.1.A.1		
4	421	CR	2 of 3 (see rubric)	6-8.PS.4.A.2		
5	421	TE	4, 3, 1, 2	6-8.ESS.1.B.1		
6	425	MS	organs, tissues,	6-8.LS.1.A.3		

Illustration: Item Separation Chart



Multi-Point Items in Workshop OIB

- Three-point items appear 3 times in the OIB: once for score point 1, once for point 2, and once for point 3.
- Ask yourself: what knowledge and skill does a student need to earn the first score point? How about the second?

Benchmarks for *Proficient*

- The benchmark is expressed as a range.
  - Based on assessment data, it is expected that your *Proficient* bookmark will be consistent with the benchmarks.
  - If you place your bookmark in this range, the percent of students classified as *Proficient* and above will be reasonably consistent with the benchmark.

Benchmarks for Grade-Level Tests

- You will see benchmarks associated with the performance of Missouri students on the National Assessment of Educational Progress (NAEP).
- Based on assessment data, your *Proficient* bookmark probably be within the benchmark range.
  - If you place your bookmark in this range, the percent of students classified as *Proficient* and above will be reasonably consistent with the benchmark.
  - It's okay to place your bookmark outside this range; but if you do, you'll need to give a content-based rationale to share with stakeholders.

About the Grade-Level Benchmarks

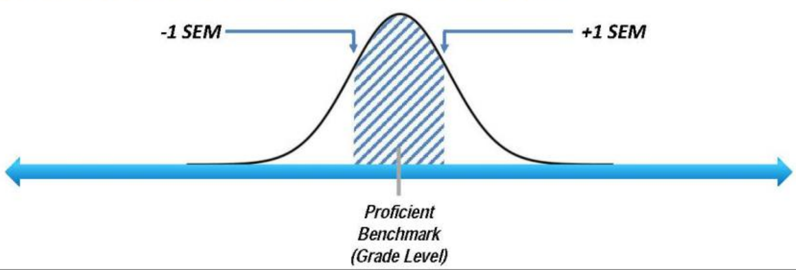
- The NAEP *Proficient* benchmark reflects the type of appropriately rigorous performance standard expected of students on the MAP.
  - However, NAEP *Proficient* is not exactly the same as MAP *Proficient.*
  - This is why you will see a benchmark *range* instead of a single page or point.
  - Brief descriptors of the NAEP levels will be provided so you can see the similarities and differences.

Benchmarks for End-of-Course Tests

- A benchmark panel of Missouri educators was convened in May of 2019
- Using a combination of NAEP and historical performance, the panel developed a recommendation for estimated performance
- BUT: the committee did not feel like any of the available external data was truly appropriate for use as a standard setting benchmark
- IN ADDITION: The committee did not feel comfortable with a single value for a recommendation, and instead a range of expected performance
- AS A RESULT: the benchmark data is provided for your information, but we do not have a reliable expectation that the committee's recommendation will be consistent with the benchmark

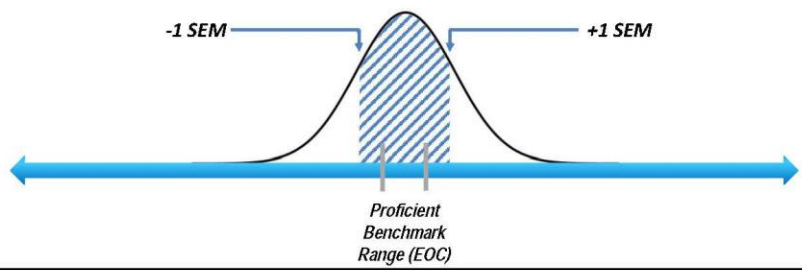
Calculating the Benchmarks

- A band of ±1 *standard error of measurement* value is calculated around each benchmark.



Calculating the Benchmarks

- A band of ±1 *standard error of measurement* value is calculated around each benchmark.



Using the Benchmarks

- Benchmarks are provided for you to focus your attention in the ordered item booklet (OIB) for the *Proficient* bookmark.
  - You should consider this benchmark range when placing your *Proficient* bookmark.
  - For the Grade-Level tests, your *Proficient* bookmark will probably fall in this range. However, your bookmark should reflect the content-based expectations of the threshold student.
  - For the EOC tests, the *Proficient* range is provided for your information and guidance, but we do NOT have a set expectation that your recommendation will be consistent with the benchmark

Sample Benchmarks for Training

- **Training:** From page 3 to 6

Item Map with Benchmarks

Missouri MAP Science Standard Setting

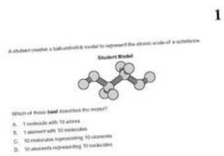
Name: ____________________

Practice Item Map

OIB Page	Location	Type	Key	Standard	What does this item measure?	What makes this item harder than previous items?
1	400	MC	A	6-8.PS.1.A.1		
2	410	CR	1 of 3 (see rubric)	6-8.PS.4.A.2		
3	412	MS	mass, increases	6-8.PS.3.A.1		
4	421	CR	2 of 3 (see rubric)	6-8.PS.4.A.2		
5	421	TE	4, 3, 1, 2	6-8.ESS.1.B.1		
6	425	MS	organs, tissues, organelles	6-8.LS.1.A.3		
7	440	MC	C	6-8.PS.4.A.1		
			3 of 3			

Benchmark
3-6

Examining an Item



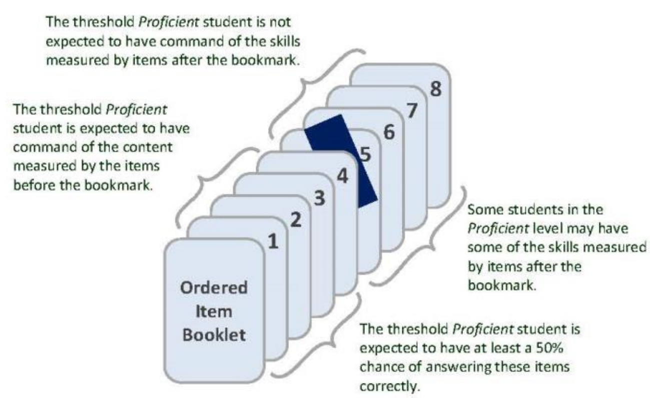
- Make a brief note to yourself about what the item measures.
  - What knowledge and skills does a student need to have in order to answer the item correctly?
  - If a student answers the item correctly, what do you know about the student?

Possible Bookmark Range

- You will find a range of items where you could set your bookmark.
  - For the *Proficient* bookmark, use the benchmarks to help identify the possible bookmark range.
  - Your possible bookmark range for *Proficient* will likely be a subset of the items in the benchmark range.
  - The possible bookmark range may be a couple of items wide, or may be more than that.
  - Do not get stuck on a single item.

Finding the Possible Bookmark Range

- Progress through the OIB until you reach an item that the threshold student would not have a 50% chance of answering correctly.
  - This is the start of your possible bookmark range.
- Keep going until you have reached the last item that a student would have a 50% chance of answering correctly.
  - The possible bookmark range ends after that page.



Recording Your Bookmark

- Place your bookmark within your Possible Bookmark Range.
    - Use the PLDs, the benchmarks, and your professional judgment as guides.
- Record the page number **after** your bookmark.

Bookmark Worksheet

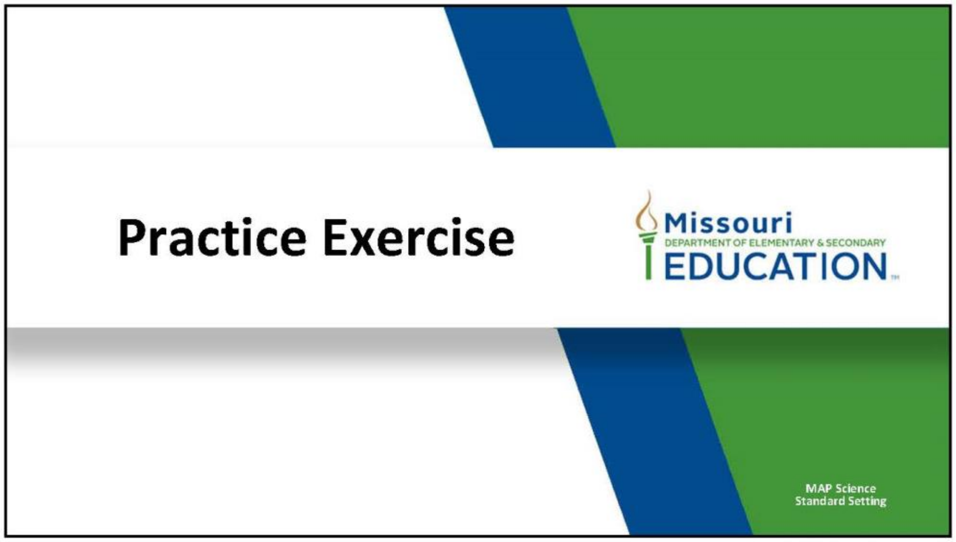
- Write your bookmarks on the *Bookmark Worksheet*.
    - You will place three bookmarks.
    - Write a content-based rationale for why you felt this was the appropriate location for your bookmark.

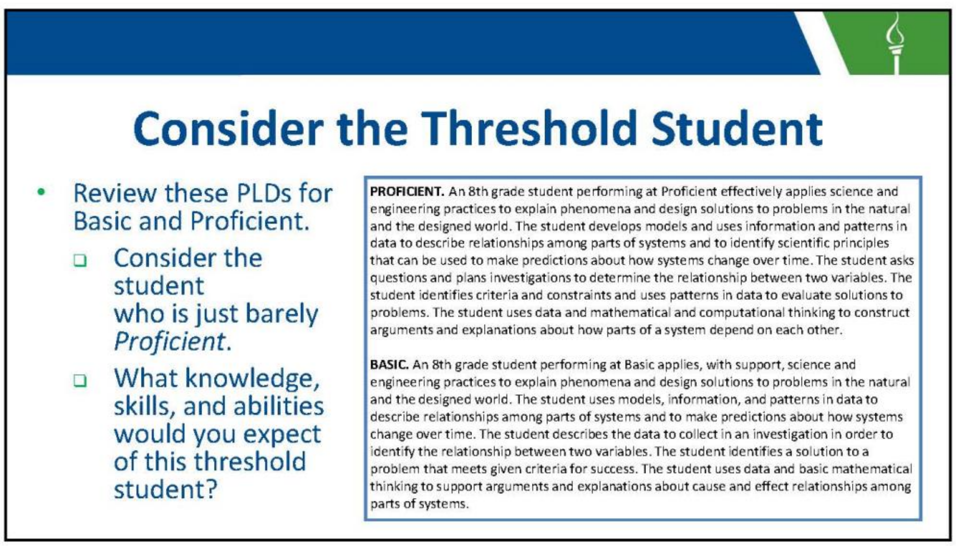
Recording Your Bookmarks

- You will record your bookmarks and your rationales using the hard copy rating forms.
- Make sure to provide a clear content-based rationale that explains why you placed your bookmark where you did.
- Complete your rating form and return to your facilitator.

Pacing

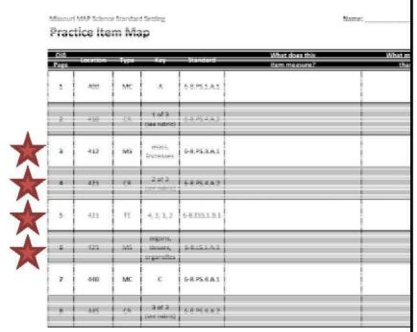
- Some people will take longer than others to study the test items and place their bookmarks.
  - During conversations, please be considerate of others at your table and in the room.
  - If you finish earlier than your neighbors, you may wish to check-in with your facilitator, leave your materials at your table, and take a short break.


Practice Exercise
Missouri
DEPARTMENT OF ELEMENTARY & SECONDARY
EDUCATION
MAP Science
Standard Setting


Consider the Threshold Student
Review these PLDs for Basic and Proficient.
Consider the student who is just barely *Proficient*.
What knowledge, skills, and abilities would you expect of this threshold student?
PROFICIENT. An 8th grade student performing at Proficient effectively applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student develops models and uses information and patterns in data to describe relationships among parts of systems and to identify scientific principles that can be used to make predictions about how systems change over time. The student asks questions and plans investigations to determine the relationship between two variables. The student identifies criteria and constraints and uses patterns in data to evaluate solutions to problems. The student uses data and mathematical and computational thinking to construct arguments and explanations about how parts of a system depend on each other.
BASIC. An 8th grade student performing at Basic applies, with support, science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses models, information, and patterns in data to describe relationships among parts of systems and to make predictions about how systems change over time. The student describes the data to collect in an investigation in order to identify the relationship between two variables. The student identifies a solution to a problem that meets given criteria for success. The student uses data and basic mathematical thinking to support arguments and explanations about cause and effect relationships among parts of systems.

Consider the Benchmarks

- The *Proficient* benchmarks for training are:
  Lower: 3
  Upper: 6
- Make a note of these benchmarks on your item map.

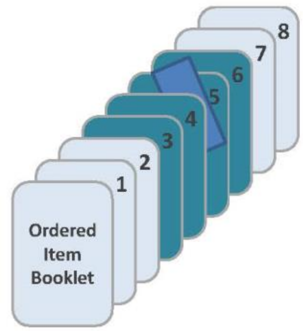


Study the Test Items

- For each question, ask yourself:
  - what does the item measure?
  - what makes the item harder than previous items?

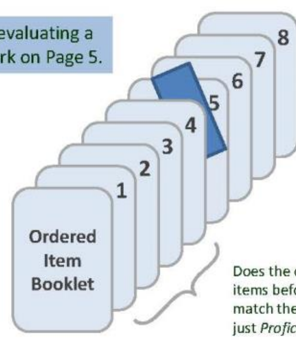
Place Your Bookmark

- Consider the *Proficient* threshold student.
- The student is expected to have at least a 50% chance of answering items correctly before the bookmark.
- The probability after the bookmark is less than 50%, but not zero.



Evaluating a Bookmark Holistically

Imagine you are evaluating a *Proficient* bookmark on Page 5.



Does the content measured by the items before the bookmark best match the content you expect of the just *Proficient* student?

Write a Rationale

- Good rationales link the content of the items in front of the bookmark to the content-based expectations for the threshold student.
  - For example, "Students must describe important parts of the cell in a real-world context, as expected of the threshold student."
  - Or, "Students have to begin to make simple inferences from empirical data, not just report those data, as listed in PLDs."
- Not-so-good rationales don't make reference to the content of the items.
  - For example, "The second score point for the four-point item is just after this bookmark."

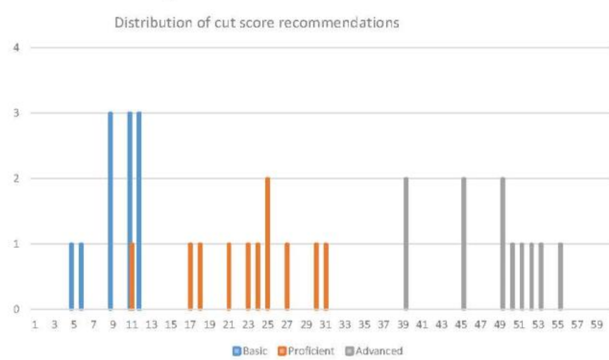
Make Your Bookmark Placements

- Write your bookmark placement on your Training Bookmark Worksheet.
- Turn your Worksheet over when you're done.

Example: Round 1 Feedback

Level	Recommended Cut
Basic	6
Proficient	22
Advanced	48

Example: Round 1 Feedback



Discussion of Round 1 Ratings

- In the actual workshop, you will discuss your Round 1 bookmarks at your table.
- Feel free to discuss:
  - Your bookmarks
  - The benchmarks for the *Proficient* cut score
  - Your possible bookmark ranges (and any overlaps)
- After discussion, you will have a second opportunity to make bookmark judgments.
  - You can change any, all, or none of your bookmarks.
  - Bookmark placement is always an individual activity.

Suggestions for Discussions

- Practice active listening.
- Be open to changing your mind.
- Work to understand your colleagues' rationales for their bookmark placements.
- In a respectful manner, feel free to ask questions of your colleagues.
- Do not discuss your bookmarks until everyone at the table has placed theirs.
- Keep the contents of your discussions private.

Example: Round 2 Feedback

Level	Recommended Cut
Basic	6
Proficient	22
Advanced	48

	Below Basic	Basic	Proficient	Advanced
Grade 5	25%	25%	25%	25%

Round 3

- After Round 2, you will discuss your bookmark placements *across tables*.
  - Again, you will share where you placed your bookmarks and why you placed them there.
  - You will also consider the *impact data* which details the estimated percentage of students that would be classified into each of the four performance categories.
- Then you will place your Round 3 bookmarks.
  - Bookmark placement is always an individual activity.

After Round 3

- After Round 3, your facilitator will show you a presentation of your Round 3 recommendations, plus the recommendations from all grades and courses.
  - You will be asked to look at the articulation of the performance standards across tests.
  - You may wish to consider adjustments to your recommendations to improve the articulation across grades.

Across-Grade Discussion

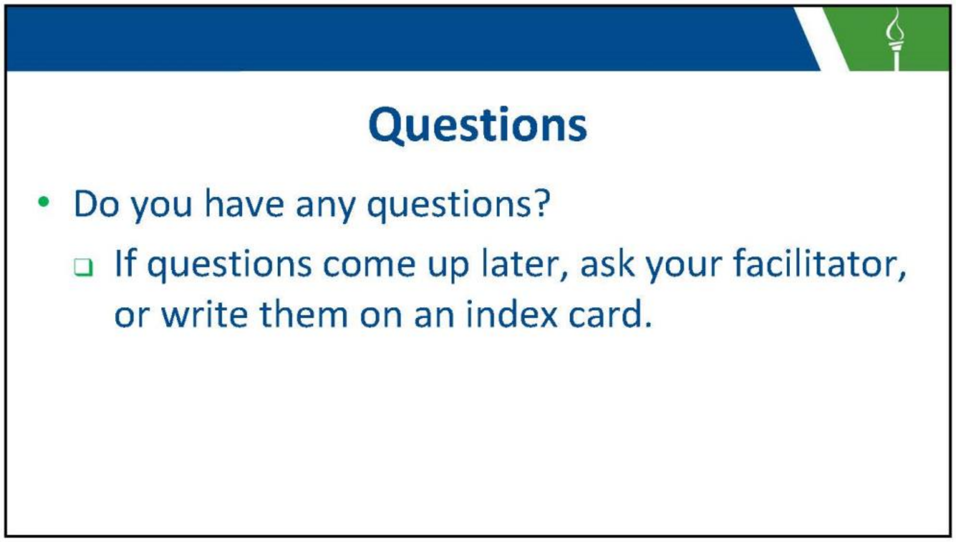
- On Thursday, table leaders will convene in an across-grade discussion to examine the recommendations.
  - If needed, table leaders will recommend adjustments to the cut scores to improve articulation across grades.
  - All participants will review the recommendations from Round 3 of the Bookmark Procedure before the across-grade discussion.

After the Workshop

- Your recommendations will be considered by DESE.
  - The recommendations from all groups will be considered by DESE and its technical advisors.

Workshop Structure

- Study PLDs
- Study OIB and make Round 1 ratings
- Discuss Round 1 at tables
- Make Round 2 ratings
- Discuss Round 2 as a group
- Make Round 3 ratings
- Review recommendations
- Across-grade discussion


Questions
• Do you have any questions?
❑ If questions come up later, ask your facilitator, or write them on an index card.

Missouri MAP Science Standard Setting

Name: ____________________

Practice Item Map

OIB Page	Location	Type	Key	Standard	What does this item measure?	What makes this item harder than previous items?
1	400	MC	A	6-8.PS.1.A.1		
2	410	CR	1 of 3 (see rubric)	6-8.PS.4.A.2		
3	412	MS	mass, increases	6-8.PS.3.A.1		
4	421	CR	2 of 3 (see rubric)	6-8.PS.4.A.2		
5	421	TE	4, 3, 1, 2	6-8.ESS.1.B.1		
6	425	MS	organs, tissues, organelles	6-8.LS.1.A.3		
7	440	MC	C	6-8.PS.4.A.1		
8	445	CR	3 of 3 (see rubric)	6-8.PS.4.A.2		

Missouri DEPARTMENT OF ELEMENTARY & SECONDARY EDUCATION

Missouri Assessment Program
Science Standard Setting

Practice
Ordered Item Booklet

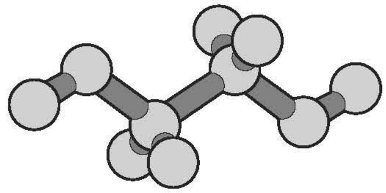
MAP Science
Standard Setting

July 2019

1

A student creates a ball-and-stick model to represent the atomic scale of a substance.

Student Model



Which of these **best** describes the model?

A. 1 molecule with 10 atoms

B. 1 element with 10 molecules

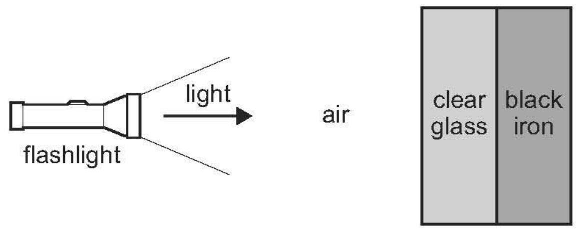
C. 10 molecules representing 10 elements

D. 10 elements representing 10 molecules

2

During an experiment, a teacher uses a flashlight and shines the light toward a pane of clear glass with a black iron backing. A diagram of the experiment is shown.

Light Experiment



Part A: Describe the path of the light as it travels from the flashlight to the air.

____

____

____

____

Part B: Describe the path of the light as it travels from the air to the pane of clear glass.

____

____

____

____

Part C: Explain why the students would **not** see the light travel through the black iron backing.

____

____

____

____

3

A student is comparing characteristics of three toy cars.

Characteristics of Three Toy Cars

Toy Car	Speed (meters/second)	Mass (kilograms)	Kinetic Energy (Joules)
1	2	1	2
2	2	2	4
3	2	4	8

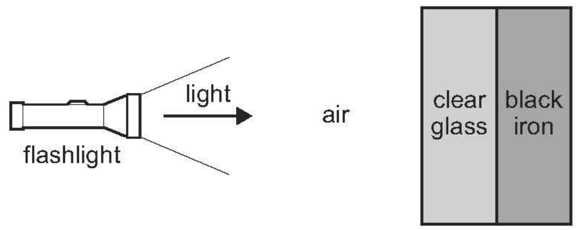
Circle a word or phrase from each set of options to complete the following sentence based on the data provided in the table.

As the (speed / mass) increases, the kinetic energy of the car (increases / decreases / stays the same).

4

During an experiment, a teacher uses a flashlight and shines the light toward a pane of clear glass with a black iron backing. A diagram of the experiment is shown.

Light Experiment



Part A: Describe the path of the light as it travels from the flashlight to the air.

Part B: Describe the path of the light as it travels from the air to the pane of clear glass.

Part C: Explain why the students would **not** see the light travel through the black iron backing.

5

The chart shows some distances between objects in the universe.

Universe Information

Objects in the Universe	Approximate Distance between Objects
Earth and the moon	382,500 kilometers
Earth and the sun	149.6 million kilometers
Neptune and the sun	4,495.1 million kilometers
Earth and the star Proxima Centauri	40,208,000 million kilometers

Identify where each measurement should be placed in the model to compare distances in the universe. Write the number from each measurement in one of the boxes in the model.

Comparing Distances in the Universe

Longest Distance — Shortest Distance

1 Distance from Earth to the sun

2 Distance from Earth to the moon

3 Distance across our solar system

4 Distance across the Milky Way galaxy

6

Circle a word in each set of options to **best** describe relationships between interacting parts of the human body.

The human body is composed of systems with interacting parts. Organ systems are made of

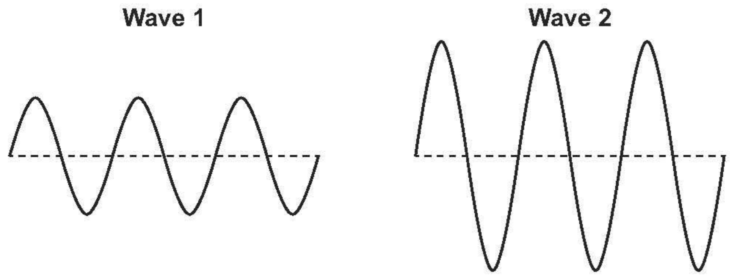
(organs / organelles / tissues), which are composed of specialized cells that work together to

form (organs / organelles / tissues). Each cell of the human body contains

(organs / organelles / tissues) with a specific function.

7

The drawings show two waves.



Which statement **best** compares these two waves?

A. Wave 1 has a higher frequency because it has a longer wavelength than wave 2.

B. Wave 1 has a higher frequency because it has a higher amplitude than wave 2.

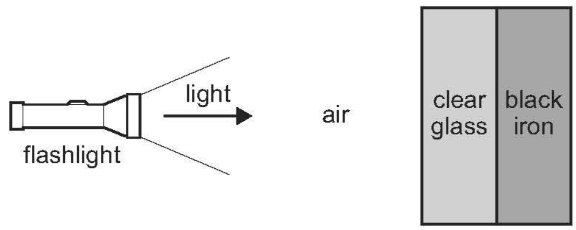
C. Wave 2 has more energy because it has a higher amplitude than wave 1.

D. Wave 2 has more energy because it has a higher frequency than wave 1.

8

During an experiment, a teacher uses a flashlight and shines the light toward a pane of clear glass with a black iron backing. A diagram of the experiment is shown.

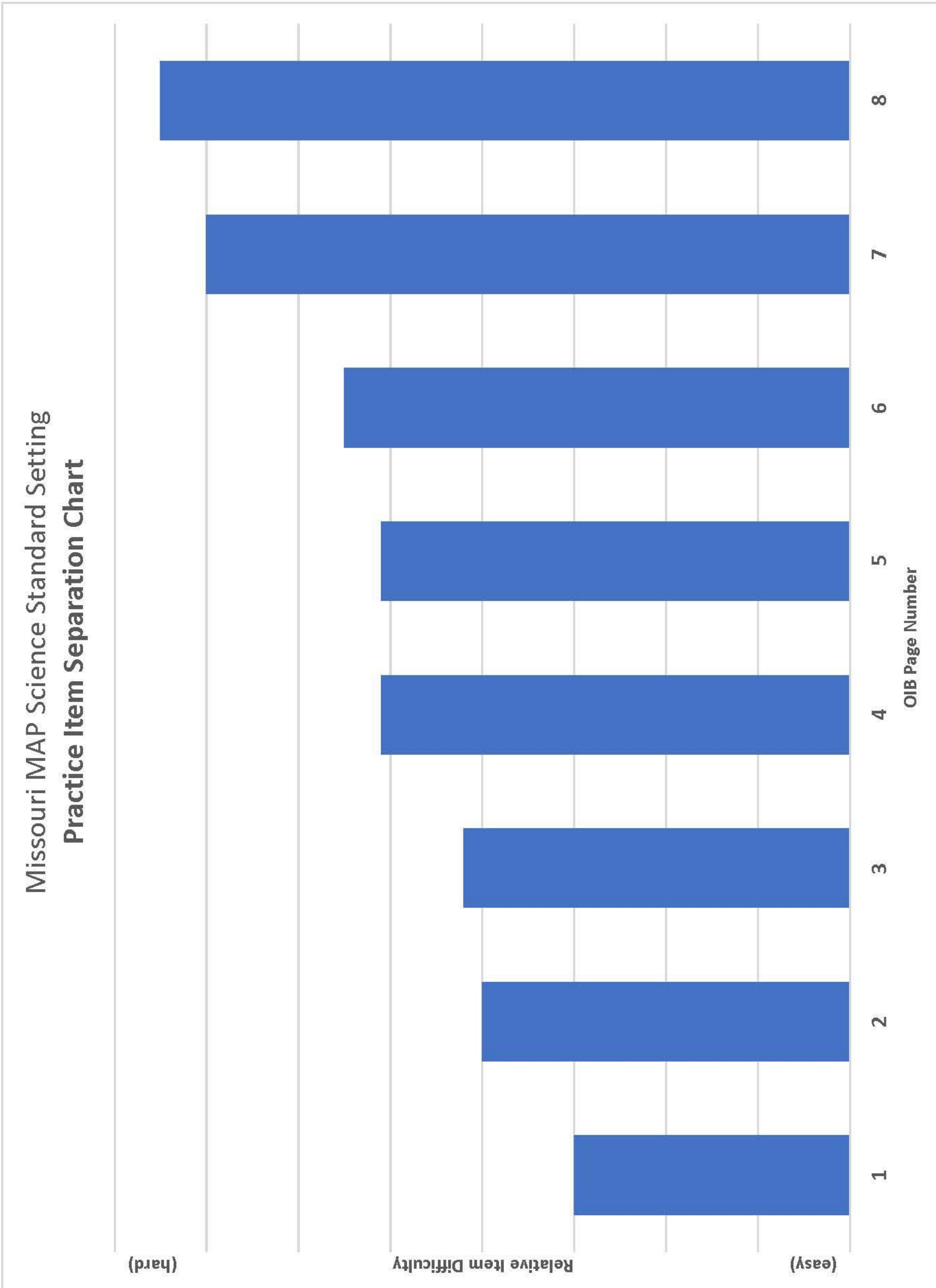
Light Experiment



Part A: Describe the path of the light as it travels from the flashlight to the air.

Part B: Describe the path of the light as it travels from the air to the pane of clear glass.

Part C: Explain why the students would **not** see the light travel through the black iron backing.


Missouri MAP Science Standard Setting
Practice Item Separation Chart
(hard)
Relative Item Difficulty
(easy)
1
2
3
4
5
6
7
8
OIB Page Number

Rubric

for OIB Pages 2, 4, and 8

Score	Description
3	This response demonstrates a thorough understanding of developing and using a model to describe that waves are reflected, absorbed, or transmitted through various materials by • describing the path of light as it travels from the flashlight to the air; • describing the path of light as it travels from the air to the pane of clear glass; and • explaining why the student would not see the light travel through the black iron backing. **The response is clear, complete, and correct.*
2	This response demonstrates a thorough understanding of two of the three key elements. **The response may contain some work that is incomplete or unclear.*
1	This response demonstrates a thorough understanding of one of the three key elements. **The response may contain some work that is incomplete or unclear.*
0	The response provides insufficient evidence to demonstrate any understanding of the concept being tested.

Exemplar Responses:

Part A (1 point)

- The light will travel in a straight line through the air.

Part B (1 point)

- The light will bend (or be refracted) when traveling into the glass.

Part C (1 point)

- The light will stop at the black iron backing.
- The light will be absorbed by the black iron backing.
    - (Note: Some reflection may occur due to luster of iron metal)

Missouri MAP Science Standard Setting – Mid-Process Evaluation – Before Round 1

The purpose of this evaluation is to obtain your feedback about the training you have received so far on the Bookmark process. Your feedback will provide a basis for determining what to review before we begin the actual process. **When you have completed the evaluation, please give it to a facilitator. *Thank you!***

	Please *consider the statements below* and mark the level of agreement or disagreement you have with each statement. Please bubble ***only one*** of the five options for each statement.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Before Round 1	1. I understood the purpose of this workshop.	O	O	O	O	O
	2. My facilitator explained things clearly.	O	O	O	O	O
	3. I understand what is meant by the threshold students.	O	O	O	O	O
	4. I understand what the ordered item booklet is.	O	O	O	O	O
	5. I understand the information presented on the item map.	O	O	O	O	O
	6. The training has given me the information I need to complete the Bookmark task.	O	O	O	O	O
	7. I understand how to make the standard setting judgments.	O	O	O	O	O
	8. I am ready to place my first bookmarks for the test.	O	O	O	O	O

If you checked "Disagree" or "Strongly Disagree" for any of the above statements, please tell us in the box below what we need to do to complete the preparation for placing the first bookmarks.

If you did *not* choose "Disagree" or "Strongly Disagree" for any of the above statements, please continue to items 9 and 10, below.

9. Before today, have you participated in a Bookmark or other standard setting workshop?

O Yes
O No

10. Please sign and date here to signal that you are ready to proceed with the standard setting process.

Signature *Date*

1. I understood the purpose of this workshop.

Response	Frequency	Percent	Mean: 4.74
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	12	26.09	
Strongly Agree	34	73.91	

2. My facilitator explained things clearly.

Response	Frequency	Percent	Mean: 4.54
Strongly Disagree	0	0.00	
Disagree	1	2.17	
Neutral	1	2.17	
Agree	16	34.78	
Strongly Agree	28	60.87	

3. I understand what is meant by the threshold students.

Response	Frequency	Percent	Mean: 4.83
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	8	17.39	
Strongly Agree	38	82.61	

4. I understand what the ordered item booklet is.

Response	Frequency	Percent	Mean: 4.85
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	15.22	
Strongly Agree	39	84.78	

5. I understand the information presented on the item map.

Response	Frequency	Percent	Mean: 4.72
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	13	28.26	
Strongly Agree	33	71.74	

6. The training has given me the information I need to complete the Bookmark task.

Response	Frequency	Percent	Mean: 4.74
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	12	26.09	
Strongly Agree	34	73.91	

7. I understand how to make the standard setting judgments.

Response	Frequency	Percent	Mean: 4.57
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	20	43.48	
Strongly Agree	26	56.52	

8. I am ready to place my first bookmarks for the test.

Response	Frequency	Percent	Mean: 4.70
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	14	30.43	
Strongly Agree	32	69.57	

9. Before today, have you participated in a Bookmark or other standard setting workshop?

Response	Frequency	Percent	Mean: 0.09
Yes	4	8.70	
No	42	91.30	

Missouri MAP Science Standard Setting – Mid-Process Evaluation – Before Round 2

The purpose of this evaluation is to gauge your readiness to continue the standard setting process now that you have completed your Round 1 bookmark judgments. Your feedback will provide a basis for determining what to review before we begin the actual process. **When you have completed the evaluation, please give it to a facilitator. *Thank you!***

	Please *consider the statements below* and mark the level of agreement or disagreement you have with each statement. Please bubble ***only one*** of the five options for each statement.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Before Round 2	1. I understood the procedure for making bookmark judgments.	O	O	O	O	O
	2. I am ready to continue with the Bookmark process.	O	O	O	O	O

If you checked "Disagree" or "Strongly Disagree" for either of the above statements, please tell us in the box below what we need to do to complete the preparation for Round 2 of bookmark placement.

If you did *not* choose "Disagree" or "Strongly Disagree" for any of the above statements, please continue to item 3, below.

3. Please sign and date here to signal that you are ready to continue with the standard setting process.

Signature *Date*

1. I understood the procedure for making bookmark judgments.

Response	Frequency	Percent	Mean: 4.74
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	26.09	
Strongly Agree	17	73.91	

2. I am ready to continue with the Bookmark process.

Response	Frequency	Percent	Mean: 4.78
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	21.74	
Strongly Agree	18	78.26	

Missouri
DEPARTMENT OF ELEMENTARY & SECONDARY
EDUCATION

Missouri Assessment Program
Science Standard Setting

July 2019

Across-Grade Discussion

- The primary goal is to examine the *vertical articulation* of the performance standards.
  - The performance standards (including the associated impact data), should be a coherent across grades and tests.
  - Performance expectations across grade levels should be consistent and be commensurate with the rigor of the assessment.

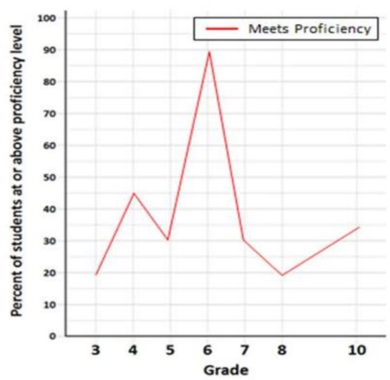
Vertical Articulation

- The primary goal is to recommend a coherent system of performance standards.
  - We are *not* trying to force things into a model that looks right but is not appropriate.
  - We *are* trying to identify discrepancies or inconsistencies that could cause confusion or distrust of the results from test users.

What is driving the motivation?

Example:
State-Level Math Proficiency Rates 2013

*Does it seem reasonable that **90% of students meet proficiency** in grade 6, but only **30% meet proficiency** in grades 5 and 7?*



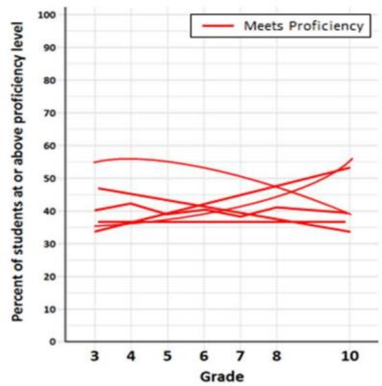
4

Vertical Articulation: What would appear to be reasonable?

Possible coherent patterns

- Linear
  - flat
  - small differences
  - increasing
  - decreasing
- Smooth Curves
  - increasing
  - decreasing

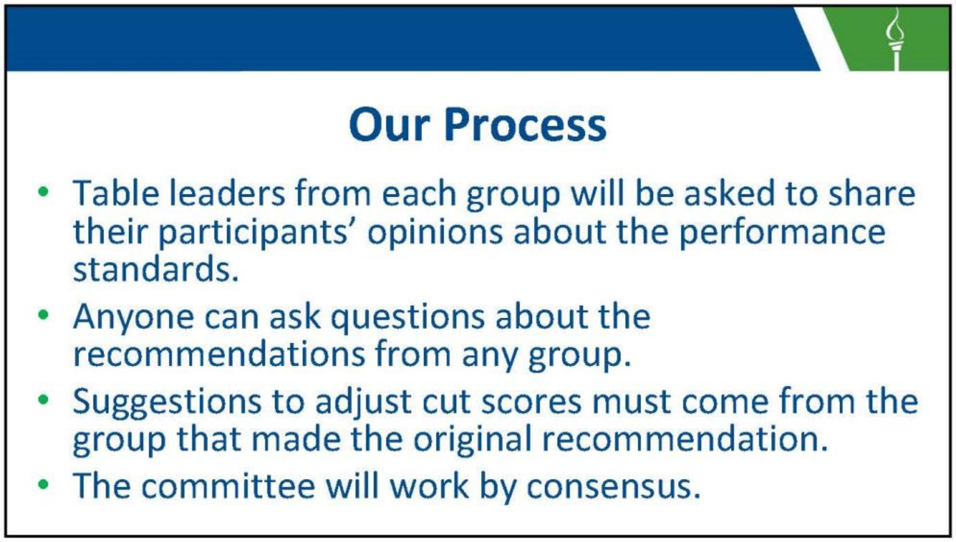
Note: These are just examples. Your judgment may differ in shape and location



5

Articulation Procedure

1. Review of PLDs across grades and content areas, looking at critical differences or shifts in content or expectations
2. Discussion of performance expectations across grades
3. Review of impact data associated with recommendations
4. Investigate possible recommendations for adjustments
5. Review resulting impact data and finalize recommendations


Our Process
Table leaders from each group will be asked to share their participants' opinions about the performance standards.
Anyone can ask questions about the recommendations from any group.
Suggestions to adjust cut scores must come from the group that made the original recommendation.
The committee will work by consensus.

 Page 87

E

Performance Level Descriptors (PLDs)

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

ADVANCED

A 5th grade student performing at Advanced effectively, consistently, and appropriately applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student evaluates models and information and revises arguments and explanations by analyzing patterns in data, cause and effect relationships, and system interactions. The student conducts investigations to collect data in order to answer questions and uses criteria and constraints to evaluate solutions to a problem. The student uses mathematical and computational thinking and scientific reasoning to analyze and interpret data in order to evaluate arguments and explanations about cause and effect relationships.

PROFICIENT

A 5th grade student performing at Proficient effectively applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student develops and uses models and information to construct arguments and explanations and to identify and describe patterns in data and system characteristics. The student asks questions that can be investigated and designs solutions to problems that meet given criteria and constraints. The student uses data and mathematical and computational thinking to construct arguments and explanations about cause and effect relationships.

BASIC

A 5th grade student performing at Basic applies, with support, science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses models and information to support arguments and explanations, to identify patterns in data, and to describe parts of systems. The student identifies the data to collect in an investigation in order to answer questions or to describe possible solutions to problems. The student uses data and basic computational thinking to support arguments and explanations about cause and effect relationships.

BELOW BASIC

A 5th grade student performing at Below Basic seldom applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student occasionally identifies models and information to identify patterns in data, and to describe relationships among parts of systems. The student infrequently recognizes trends in the data collected during an investigation in order to answer questions or to identify possible solutions to problems. The student occasionally uses data and basic computational thinking to explain the cause and effect relationships.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Physical Science	**Below Basic** A student who has reached the level of *Below Basic* is able to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Matter and Its Interactions	Recognize - a phase change of water as a result of temperature change. - changes in matter such as weight and temperature during changes in substance. Identify - changes in the state of matter as a result of temperature change. - or recognize that matter is made of particles too small to be seen. - or observe properties of materials. - whether the mixing of substances produces a new substance.	Investigate - a phase change of water as a result of temperature change. - properties of materials. - to collect measurements of matter such as weight and temperature during changes in substance. - whether the mixing of substances produces a new substance. Describe evidence for - changes in the state of matter as a result of temperature change. Recognize a model that describes - matter is made of particles too small to be seen.	Investigate - and predict a phase change of water as a result of temperature change. - and describe whether the mixing of substances produces a new substance. Use evidence - to predict changes in the state of matter as a result of temperature change. - develop a model to describe that matter is made of particles too small to be seen. - of measurements to identify materials by their properties. - to support that matter is conserved during changes in substance.	Analyze and evaluate evidence - from an investigation to explain a phase change of water as a result of temperature change. - to predict changes in state of matter as a result of temperature change. - of measurement data to identify materials based upon their properties. - from an investigation for whether the mixing of substances produces a new substance. - to construct an argument that matter is conserved during changes in substance. Evaluate models to - explain different types of matter made of particles too small to be seen.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Motion Stability: Forces and Interactions	Observe or describe - an object's motion. Identify - effects of balanced and unbalanced forces on an object's motion. - electric or magnetic interactions between two objects. - forces on an object moving on different surfaces. Recognize - the gravitational force exerted by Earth on objects is directed toward the planet's center. - that force or mass affect motion.	Investigate - an object's motion. - effects of balanced or unbalanced forces on an object's motion. - forces on an object moving on different surfaces. - to describe how force or mass affects motion. Describe evidence of - electric or magnetic interactions between two objects. Identify evidence to support - the gravitational force exerted by Earth on objects is directed toward the planet's center.	Investigate - an object's motion to predict future motion. - and describe whether balanced or unbalanced forces change an object's motion. - and predict forces on an object moving on different surfaces. - to explain how force or mass affects motion. Use evidence to - support the gravitational force exerted by Earth on objects is directed toward the planet's center. - explain the electric or magnetic interaction between two objects.	Analyze models from an investigation - for an object's motion using evidence to predict future motion. - to provide evidence of balanced and unbalanced forces on an object's motion. - to explain the electric or magnetic interaction between two objects. - to explain forces on an object moving on different surface. - to predict how force or mass affects motion. Evaluate evidence to support - the gravitational force exerted by Earth on objects is directed toward the planet's center.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Energy	Identify - a relationship between speed and energy of an object. - a device that converts energy from one form to another. - a simple machine. - that energy in food comes from the sun. Recognize - energy transformation.	Identify evidence to - describe energy transformation. - describe a relationship between speed and energy of an object. Recognize a model - of a simple machine. - showing energy in food comes from the sun. - of a device that converts energy from one form to another.	Use evidence to explain - energy transformation. - a relationship between speed and energy of an object. Use a model to describe - simple machines. - energy in food comes from the sun. - design, test, or refine a device that converts energy from one form to another.	Analyze and evaluate evidence - of an explanation for a relationship between speed and energy of an object. - to predict energy transformation Evaluate a model - to explain the relationship between simple machines and force. - to explain the relationship between simple machines and force. - or refine a model of a device that converts energy from one form to another.
Waves and Their Applications in Technologies for Information	Identify or recognize - wave properties. - objects can be seen only when light is reflected or when they produce their own light.	Identify a model - of wave properties. - showing objects can be seen only when light is reflected or when they produce their own light.	Use a model to - describe wave properties. - show objects can be seen only when light is reflected or when they produce their own light.	Analyze or evaluate a model to - explain wave properties. - explain why objects can be seen only when light is reflected or when they produce their own light.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Life Science	**Below Basic** A student who has reached the highest level of the *Below Basic* level is able to complete the following:	**Basic** A student who has *just* reached the level of *Basic* is able to complete the following:	**Proficient** A student who has *just* reached the level of *Proficient* is able to complete the following:	**Advanced** A student who has *just* reached the level of *Advanced* is able to complete the following:
From Molecules to Organisms: Structure and Processes	Recognize - structures for support, survival, growth, behavior, and plant reproduction. - life cycles of plants and animals. - that plants primarily need air and water to grow. - animals respond based on information through their senses. Identify - similarities and differences between body systems.	Identify evidence - structures for support, survival, growth, behavior, and plant reproduction. - that plants primarily need air and water to grow. - similarities and differences between body systems. - life cycles of plants and animals. - how animals respond based on information through their senses.	Use evidence to - describe structures for support, survival, growth, behavior, and plant reproduction. - describe similarities and differences between body systems. - support an argument that plants primarily need air and water to grow. Use a model to describe - life cycles of plants and animals. - how animals respond based on information through their senses.	Analyze evidence to support arguments - that plants and animals have structures for support, survival, growth, behavior, and plant reproduction. - to evaluate similarities and differences between body systems. Evaluate a model to - explain life cycles of plants and animals. - explain how animals respond based on information through their senses. - support the argument that plants primarily need air and water to grow.
Ecosystems: Interactions, Energy, and Dynamics	Identify - how matter moves through organisms within an ecosystem.	Identify a model - how matter moves through organisms within an ecosystem.	Develop or use a model to - describe how matter moves through organisms within an ecosystem.	Evaluate a model to - explain how matter moves through organisms within an ecosystem.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Heredity: Inheritance and Variation of Traits	Recognize - a solution to a problem with plants and animals caused when the environment changes. Identify - characteristics inherited from parents or influenced by the environment. - variations in a species may increase survival or reproduction. - that some organisms have adaptations to survive better in an ecosystem.	Recognize evidence - of variations in a species may increase survival or reproduction. Identify evidence to describe - characteristics inherited from parents or influenced by the environment. - that some organisms have adaptations to survive better in an ecosystem. - a solution to a problem with plants and animals caused when the environment changes.	Support a claim using evidence to explain - characteristics inherited from parents or influenced by the environment. Use evidence to - explain that some organisms have adaptations to survive better in an ecosystem. - explain a solution to a problem with plants and animals caused when the environment changes. - describe variations in a species may increase survival or reproduction.	Analyze or evaluate evidence to - explain variations in a species may increase survival or reproduction. - construct an explanation for a solution to a problem with plants and animals caused when the environment changes. - to argue that some characteristics are inherited from parents or influenced by the environment. Use models and evidence to argue - that some organisms have adaptations to survive better in an ecosystem.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Earth and Space Science	**Below Basic** A student who has reached the highest level of the *Below Basic* level is able to complete the following:	**Basic** A student who has *just* reached the level of *Basic* is able to complete the following:	**Proficient** A student who has *just* reached the level of *Proficient* is able to complete the following:	**Advanced** A student who has *just* reached the level of *Advanced* is able to complete the following:
Earth's Place in the Universe	Recognize - relationships between amount of daylight and time of year. - change in landscape over time. - daily patterns of shadows and seasonal changes in the night sky. Identify - differences in brightness among stars in the sky and the Sun.	Observe and describe - relationships between amount of daylight and time of year. - daily patterns of shadows and seasonal changes in the night sky. Identify evidence for - causes for change in landscape over time. - differences in brightness among stars in the sky and the Sun.	Use evidence to - support explanations of the relationships between daylight and time of year. - describe the difference in brightness of the Sun compared to other stars is due to distance. Use a model to - describe evidence for changes in landscape over time. - reveal observable daily patterns of shadows and seasonal changes in the sky.	Analyze or evaluate evidence to - support a claim explaining the relationships between daylight and time of year. Use a model and evidence to - explain changes in landscape over time. - argue that the difference in brightness of the Sun compared to other stars is due to distance. - explain observable daily patterns of shadows and seasonal changes in the sky.
Earth's Systems	Recognize - how natural processes shape Earth's surfaces. - typical weather conditions. Identify - patterns in Earth's features. - reservoirs of water on Earth. - climates in different regions. Describe - the ways in which the four Earth spheres interact.	Investigate - how natural processes shape Earth's surface. - and describe reservoirs of water on Earth. Identify a model to - describe the ways in which the four Earth spheres interact. - represent data of typical weather conditions. - recognize patterns found in data to describe Earth's features. Describe evidence of - climates in different regions.	Investigate and provide evidence for - how natural processes shape Earth's surface. Develop or use a model to describe - the ways in which the four Earth spheres interact. - patterns in data of Earth's features. - typical weather conditions. - differences in water distribution on Earth. Evaluate evidence of - climates in different regions.	Use evidence from an investigation to evaluate - how natural processes shape Earth's surface. Analyze or evaluate models to explain - multiple ways in which the four Earth spheres interact. - patterns of Earth's features. - differences in water distribution on Earth. - typical weather conditions and predict future conditions. Evaluate evidence to describe and predict - climates in different regions.

GRADE 5 RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Earth and Human Activity	Recognize - solutions to reduce the impacts of natural Earth processes on humans. Identify - a solution that reduces the impacts of weather-related hazards. - a way to protect the environment.	Describe - solutions to reduce the impacts of natural Earth processes on humans. - evidence for a way to protect the environment. Recognize a claim - to a solution that reduces the impacts of weather-related hazards.	Use evidence to explain - ways to protect the environment. Develop or compare a claim - to a solution that reduces the impacts of weather-related hazards. - of solutions to reduce the impacts of natural Earth processes on humans.	Analyze or evaluate evidence for - solutions to reduce the impacts of natural Earth processes on humans. - a claim to a solution that reduces the impacts of weather-related hazards. Use evidence to combine information about and argue for - ways to protect the environment.

Engineering and Technology Science	**Below Basic** A student who has reached the highest level of the *Below Basic* level is able to complete the following:	**Basic** A student who has *just* reached the level of *Basic* is able to complete the following:	**Proficient** A student who has *just* reached the level of *Proficient* is able to complete the following:	**Advanced** A student who has *just* reached the level of *Advanced* is able to complete the following:
Engineering Design	Identify design constraints and criteria.	Describe the design constraints and criteria.	Define a simple design problem, including constraints and criteria.	Explain a simple design problem, including constraints and criteria.
	Recognize a possible solution to an engineering problem.	Describe a possible solution to an engineering problem.	Generate and compare multiple possible solutions to an engineering design problem.	Use several sources to generate and compare multiple possible solutions to an engineering problem.
	Recognize ways to improve a model or prototype.	Carry out tests to improve a model or prototype.	Carry out tests to improve a model or prototype by controlling variables.	Carry out tests and analyze data to improve a model or prototype by controlling variables and identifying failures.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

ADVANCED

An 8th grade student performing at Advanced effectively, consistently, and appropriately applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student evaluates how well models, information, and patterns in data describe relationships among parts of systems, and uses scientific principles and reasoning to make predictions about how systems change over time. The student plans and evaluates investigations designed to determine the relationship between two variables. The student uses patterns in data to determine which solution to a problem best meets the criteria for success. The student uses data, mathematical and computational thinking, and scientific reasoning to construct and evaluate arguments and explanations about how parts of a system depend on each other.

PROFICIENT

An 8th grade student performing at Proficient effectively applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student develops models and uses information and patterns in data to describe relationships among parts of systems and to identify scientific principles that can be used to make predictions about how systems change over time. The student asks questions and plans investigations to determine the relationship between two variables. The student identifies criteria and constraints and uses patterns in data to evaluate solutions to problems. The student uses data and mathematical and computational thinking to construct arguments and explanations about how parts of a system depend on each other.

BASIC

An 8th grade student performing at Basic applies, with support, science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses models, information, and patterns in data to describe relationships among parts of systems and to make predictions about how systems change over time. The student describes the data to collect in an investigation in order to identify the relationship between two variables. The student identifies a solution to a problem that meets given criteria for success. The student uses data and basic mathematical thinking to support arguments and explanations about cause and effect relationships among parts of systems.

BELOW BASIC

An 8th grade student performing at Below Basic seldom applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student occasionally identifies models, information, and patterns in data to describe relationships among parts of systems and to make predictions about how systems change over time. The student infrequently recognizes trends in the data collected during an investigation in order to identify the relationship between two variables. The student can sometimes identify a solution to a problem. The student occasionally uses data and basic mathematical thinking to explain the cause and effect relationships among parts of systems.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Physical Science	**Below Basic** A student who has reached the level of *Below Basic* is able to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Matter and Its Interactions	Identify - atomic or molecular structures. - the occurrence of chemical reactions. - different factors that cause changes in thermal energy. Organize - information about the chemical properties of substance. Describe - the effects of temperature changes on substances. - how mass is conserved in a chemical reaction.	Identify a model to describe - atomic or molecular structures. - changes to substances caused by temperature changes. - which describes how mass is conserved in a chemical reaction. - a device that uses changes in thermal energy. Identify evidence - collected information about the properties of designed materials as they relate to their chemical properties. Use data to - identify the occurrence of chemical reactions.	Use models to describe - atomic composition of molecular structures. - the physical changes that occur when the materials thermal energy changes. - how mass is conserved in chemical reactions. Communicate evidence to describe - the occurrence of chemical changes. - a design and modify a device that uses changes in thermal energy by chemical processes. Collect and synthesize data - information about the chemical properties of designed materials.	Analyze and/or evaluate models - to describe atomic composition of molecular structures. - to explain the changes in particle motion, temperature, and state of matter that occur when the materials thermal energy changes. - which use the total number of atoms to explain how mass is conserved in chemical reactions. Analyze or evaluate evidence to - explain the properties of substances for the occurrence of chemical changes. - collect, test, and optimize a device that uses changes in thermal energy by chemical processes. Analyze and interpret data - about the chemical properties of designed materials to evaluate potential impact.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Motion Stability: Forces and Interactions	Identify or recognize - a solution which minimizes the force of an object during a collision. - that electric or magnetic forces are a noncontact force between objects of various masses. - data for the change in an object's motion on the forces on the object and the mass of the object. - the effects of electric and magnetic fields on objects. Organize - provided evidence of gravitational interaction between objects of various masses or distances.	Identify evidence - to describe solutions which minimize the force of an object during a collision. - from a graph of gravitational interaction between objects of various masses or distances. - to explain the effects of electric or magnetic fields between objects of various masses. - and collect evidence for the effects of electric and magnetic fields on objects. Investigate - the change in an object's motion on the forces on the object and the mass of the object.	Conduct an investigation - into the change in an object's motion on the forces on the object and the mass of the object. - regarding the force exerted by electrical and magnetic fields. - design a solution which minimizes the force of an object during a collision. Use a model to - provide evidence of the gravitational interaction between objects of various masses or distances. - communicate evidence which explains the effects of electric and magnetic fields on objects.	Plan and conduct an investigation - design and evaluate a solution which minimizes the force of an object during a collision. - to evaluate the change in an object's motion on the sum of forces on the object and the mass of the object. - regarding the force exerted by electrical and magnetic fields. Analyze and evaluate models to - explain the effects of electric and magnetic fields on objects. - argue for the gravitational interaction between objects of various masses and distances.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Energy	Describe - the relationship of kinetic energy to the mass and speed of objects. - how the temperature of objects depends upon energy, mass, and types of matter. - how thermal energy is transferred in a device. Identify - potential energy in different systems. - opportunities when kinetic energy of an object changes.	Organize data to describe - the relationship of kinetic energy to the mass and speed of objects. Identify a model to describe - the interactions of objects in a system based upon potential energy. - changes in temperature relating to energy transfer, mass, and types of matter. - a device which minimizes or maximizes thermal energy transfer. - evidence that supports the kinetic energy of an object changes, energy is transferred.	Analyze a model to describe - the relationship of kinetic energy to the mass or speed of objects. - changes in temperature relating to energy transfer, mass, and types of matter. - an argument that when the kinetic energy of an object changes, energy is transferred. Use models to - explain the interactions of objects in a system based upon different forms of potential energy. - design a device that either minimizes or maximizes thermal energy transfer.	Generate, collect, and interpret models - to explain the relationship of kinetic energy to the mass and speed of objects. - to test a device that either minimizes or maximizes thermal energy transfer. - explain the interactions of objects at different distances in a system based upon different forms of potential energy. Plan and conduct an investigation to - analyze changes in temperature relating to energy transfer, mass, and types of matter. - construct and use an argument that when the kinetic energy of an object changes, energy is transferred.
Waves and Their Applications in Technologies for	Identify or observe - various wave properties and behavior. - how waves interact with different media.	Identify models to describe - wave properties and behavior. - wave interactions with different media.	Use models to describe - wave properties and behavior. - wave interactions with different media.	Develop, analyze or evaluate models - to explain wave properties and behavior. - that can be used to collect data which describes wave interactions with different media.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Life Science	**Below Basic** A student who has reached the level of *Below Basic* is able to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
From Molecules to Organisms: Structure and Process	Identify or recognize - that living things are made of cells. - body systems that interact to carry out key body functions. - the effects of resource availability on individual organisms and populations of organisms in an ecosystem. - abiotic and biotic factors in an ecosystem. Describe - how cells or parts of cells work together. - how interacting groups of cells perform life functions.	Use data from investigations as evidence - that things are made of cells. - for the effects of resource availability on individual organisms and populations of organisms in an ecosystem. Identify models which describe - how cells or parts of cells work together. - the pattern between abiotic and biotic factors. Support an argument - for how interacting groups of cells perform life functions. Recognize evidence of - how body systems interact to carry out key body functions.	Use models and data from investigations as evidence - that living things are made of cells. - for the effects of resource availability on individual organisms and populations of organisms in an ecosystem Develop models to describe - the function of cells or how parts of cells work together. - a generalized hypotheses about interaction patterns among biotic or abiotic factors in ecosystems. Critique an argument of - how interacting groups of cells perform life functions. Identify evidence - that body systems interact to carry out key body functions.	Use models and data from investigations as evidence - that living things are made of cells and distinguish between living and non-living things. - to evaluate the effects of resource availability on individual organisms and populations of organisms in an ecosystem. Analyze or evaluate models - to support an argument for how cells or parts of cells work together. - of hypotheses about interaction patterns among biotic and abiotic factors in ecosystems. Develop an argument of how - interacting groups of cells perform life functions. Evaluate evidence - that body systems interact to carry out key body functions.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Ecosystems: Interactions, Energy, and Dynamics	Identify - animal behaviors and plant structures related to reproduction. - different interactions of organisms in ecosystems. - examples of photosynthesis and cellular respiration. Describe - how organisms within an ecosystem depend upon living and nonliving components. - the effects of human action upon biodiversity. - how energy is used in organisms..	Identify a model of - chemical reactions involving food molecules to explain how energy is used in organisms. Construct an explanation of - the role of photosynthesis or cellular respiration in the cycling of matter and flow of energy into and out of organisms. - how organisms within an ecosystem depend upon the cycling of living and nonliving components. - the interaction patterns among organisms in ecosystems. - how animal behaviors and plant structures are related to reproduction. Evaluate solutions - that minimize the effects of human actions upon biodiversity.	Develop a model which describes - the cycling of matter and flow of energy among living and nonliving parts of an ecosystem. Gather and use evidence - to support the claim that animal behaviors or plant structures affect reproduction. - construct an argument how physical or biological components of an ecosystem affect populations. Construct an explanation based on evidence of - how genetic and environmental factors affect organisms. - the role of photosynthesis and cellular respiration in the cycling of matter and flow of energy into and out of organisms. Evaluate the benefits and limitations of - a design maintaining an ecosystem.	Develop and evaluate a model which describes - the cycling of matter and flow of energy among living and nonliving parts of an ecosystem. Gather and use evidence to construct an explanation - that animal behaviors or plant structures affect reproduction. Evaluate a scientific explanation based on evidence to explain - how genetic and environmental factors affect organisms. - the role of photosynthesis and cellular respiration in the cycling of matter and flow of energy into and out of organisms. Evaluate an argument of - how physical or biological components of an ecosystem affect populations. Evaluate the benefits and limitations of - two designs for maintaining an ecosystem.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Biological Evolution: Unit and Diversity	Identify - patterns of relatedness of organisms and fossils based on anatomy. - humans influence the inheritance of desired traits in organisms. Predict - changes in traits within populations over time. Recognize - that specific traits will lead to increases or decreases in survival or reproduction chances.	Identify evidence to explain - patterns of relatedness of organisms and fossils based on anatomy. - why specific traits will lead to increases or decreases in survival or reproduction chances. Identify - technologies that have changed the way humans influence the inheritance of desired traits in organisms. Use mathematical relationships (models) to explain - changes in traits within populations over time.	Interpret data to explain - patterns of relatedness of organisms and fossils based on anatomy. Construct an explanation to explain - why specific traits will lead to increases or decreases in survival or reproduction chances. Gather and synthesize information - about technologies that have changed the way humans influence the inheritance of desired traits in organisms. Interpret models to support how - natural selection may lead to increases and decreases of specific traits in populations over time.	Analyze and evaluate data to - explain patterns of relatedness of organisms and fossils based on anatomy. Evaluate - an explanation used to explain why specific traits will lead to increases or decreases in survival or reproduction chances. - evidence of technologies that have changed the way humans influence the inheritance of desired traits in organisms. - a model to support how natural selection may lead to increases and decreases of specific traits in populations over time.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Earth and Space Science	Below Basic A student who has reached the level of *Below Basic* is able to successfully address some, but not all, of the following:	Basic A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	Proficient A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	Advanced A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Earth's Place in the Universe	Identify - patterns involving the Sun and the Moon based upon their relative positions. - properties of objects in the solar system. Recognize - how gravity affects motion within the solar system and within galaxies. Use rock strata to - organize Earth's history.	Identify a model to explain - patterns involving the Sun and the Moon based upon their relative positions. - how gravity explains motion within the solar system and within galaxies - describe the properties of objects in the solar system. Identify a scientific explanation - from rock strata to organize Earth's history.	Use a model to explain - patterns and make predictions involving the Sun and the Moon based upon their relative positions. - how gravity explains motion within the solar system and within galaxies. Construct a scientific explanation - from rock strata to organize Earth's history. Analyze data to identify - patterns in the properties of objects in the solar system.	Evaluate models to explain - patterns and make predictions involving the Sun and the Moon based upon their relative positions. Gather and analyze information - to develop a model of how gravity explains motion within the solar system and within galaxies. - to explain the difference in the properties of objects in the solar system. Evaluate a scientific explanation - from rock strata to organize Earth's history.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Earth's Systems	Describe - the role of energy in the cycling of Earth's materials. - the roles of energy and gravity in the water cycles. - how Earth's processes are related to the distribution of natural resources. - evidence of past tectonic plate motions. - Earth's history using rock formations and fossils. - how heat and Earth's rotation produce differences in atmospheric and oceanic circulation patterns that lead to different climates. Identify - geological processes that create geological features. Relate - the interaction of air masses to changes in weather.	Identify a model to describe - the role of energy in the cycling of Earth's materials. - the roles of energy and gravity in the water cycle. - how heat and Earth's rotation produce differences in atmospheric and oceanic circulation patterns that lead to different climates. Identify evidence - from Earth's processes to explain the distribution of natural resources. - from rock formations and fossils to explain Earth's history. - to explain how geological processes of different time and spatial scales create geological features. - of the interaction of air masses to explain changes in weather. Organize data that provide evidence - of past tectonic plate motions.	Identify evidence to explain - how geological processes of varying time and spatial scales create geological features. Develop and use a model to explain - the roles of energy and gravity in the water cycle. - the role of energy in the cycling of Earth's materials. - how heat and Earth's rotation produce differences in atmospheric and oceanic circulation patterns that lead to different climates. Construct a scientific explanation based on evidence that explains - the distribution of natural resources. Analyze information - from rock formations and fossil evidence to explain Earth's history. - and interpret data of past tectonic plate motions. - and evidence of the interaction of air masses to explain changes in weather.	Evaluate models and use evidence to explain - the role of energy in the cycling of Earth's materials. - the roles of energy and gravity in the water cycle. - the limitations of a model to describe how heat and Earth's rotation produce differences in atmospheric and oceanic circulation patterns that lead to different climates. Evaluate explanations - from Earth's processes to explain the distribution of natural resources. - and data of past tectonic plate motions for its usefulness in developing a model. Gather evidence - to explain how geological processes of varying time and spatial scales create geological features. - and evaluate data of the interaction of air masses to explain changes in weather. - to synthesize information from rock formations and fossil evidence to explain Earth's history.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

	Recognize - characteristics of natural hazards. Describe - human impacts on the environment. - how population growth increases the use of natural resources and causes environmental changes. Identify - factors that cause a change in global temperatures for the past century.	Design a method to - monitor or minimize human impacts on the environment. Use provided evidence to - argue that population growth increases the use of natural resources and causes environmental changes. - identify factors that cause a change in global temperatures for the past century. - identify patterns about natural hazards.	Design and refine a method to - monitor or minimize human impacts on the environment. Analyze evidence - that population growth increases the use of natural resources and causes environmental changes. - of factors that cause a change in global temperatures for the past century. - related to strategies to minimize dangers from natural hazards through forecasting and technology.	Evaluate competing designs to - monitor or minimize human impacts on the environment. Evaluate data for use as evidence - that population growth increases the use of natural resources and causes environmental changes. - of factors that cause a change in global temperatures for the past century. - related to strategies to minimize dangers from natural hazards through forecasting and technology which mitigates effects.

GRADE 8 PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Engineering and Technology Science	**Below Basic** A student who has reached the level of *Below Basic* is able to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Engineering Design	Describe - potential impacts of design. - how to improve a design through repeated testing. Compare - competing designs to solve a specific problem. - design solutions using test data.	Identify - the potential impacts of a design in order to define criteria and constraints. - between competing designs to solve a specific problem using criteria and constraints. Develop a model to - optimize a design through repeated testing. Analyze test data - to compare design solutions.	Analyze - the potential impacts of a design in order to prioritize criteria and constraints. - test data to support an argument for an optimal design. Synthesize data to develop a model to - optimize a design through repeated testing. Support an argument for - a design to solve a specific problem using criteria and constraints.	Evaluate - the potential impacts of a design in order to prioritize criteria and constraints. - test data to best support an argument for an optimal design. - data used to develop a model to optimize a design through repeated testing. Support an argument - for the best design to solve a specific problem using criteria and constraints.

PHYSICAL SCIENCE RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

ADVANCED

A student performing at Advanced effectively, consistently, and appropriately applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses information to evaluate patterns in data and revise models that support scientific claims, explain relationships among variables, and predict, based on scientific principles and reasoning, how the variables will change over time. The student revises the design of investigations in order to collect data that can describe quantitative relationships among variables. The student analyzes patterns in data to determine which solution best meets the criteria and constraints of a problem. The student uses data, mathematical and computational thinking, and scientific principles to construct explanations of scientific processes and arguments about stability and change within systems.

PROFICIENT

A student performing at Proficient effectively applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student develops models and uses information and patterns in data to support scientific arguments, describe relationships among variables, and predict how the variables will change over time. The student plans investigations to determine proportional relationships among variables. The student analyzes patterns in data to evaluate how well a solution meets the criteria and constraints of a problem. The student uses data, mathematical and computational thinking, and scientific principles to construct explanations of scientific processes and arguments about how systems and system parts will change over time.

BASIC

A student performing at Basic applies, with support, science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses models, information, and patterns in data to support scientific arguments, identify the relationship between two variables, and make predictions about how changes to one variable will affect other variables. The student describes the data to collect in an investigation in order to identify proportional relationships among variables. The student uses patterns in data to identify a solution that meets given criteria and constraints of a problem. The student uses data, basic algebraic thinking, and scientific principles to support explanations of scientific processes and arguments about how systems and system parts will change over time.

BELOW BASIC

A student performing at Below Basic seldom applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student occasionally identifies models, information, and patterns in data to describe relationships between two variables and make predictions about how changes to one variable will affect other variables. The student infrequently recognizes trends in the data collected during an investigation in order to identify the relationships among variables. The student can sometimes use patterns in data to identify a solution to a problem. The student occasionally uses data and basic algebraic thinking to explain how systems and system parts change over time.

PHYSICAL SCIENCE RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Physical Science	**Below Basic** A student who has reached the level of *Below Basic* level is to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Matter and Its Interactions	Identify - chemical and atomic properties. - an explanation for products of a simple chemical reaction. - the chemical properties that can change during a chemical reaction. - changes in chemical reaction rates. - a refined design of a chemical system. Recognize - different bulk properties of matter and its physical changes. - molecular properties of designed materials. - that mass is conserved during chemical reactions. - changes in the composition of the nucleus of the atom or the release of energy during fission, fusion, or radioactive decay.	Describe - molecular properties of designed materials. - chemical properties that can change during a chemical reaction. - changes in chemical reaction rates. - a refined design of a chemical system. - mass is conserved during chemical reactions. - changes in the composition of the nucleus of the atom or the release of energy during fission, fusion, or radioactive decay. Identify, measure, or record - different bulk properties of matter and its physical changes. - the pattern of chemical and atomic properties using the periodic table. Revise an explanation - for products of a simple chemical reaction.	Use a model to - predict chemical and atomic properties using the periodic table. - to support an argument for the conservation of mass in a chemical reaction. - illustrate changes in the composition of the nucleus of the atom and the release of energy during fission, fusion, or radioactive decay. - refine a design of a chemical system by specifying a change in conditions that would alter the amount of products at equilibrium. Use evidence - of chemical properties to explain the outcome of a chemical reaction. - to explain changes to chemical reaction rates. - through investigations to compare different bulk properties of matter and its physical changes. Construct an explanation - for products of a simple chemical reaction. - to communicate the function of a designed material based on its molecular properties. -	Evaluate models to - explain chemical and atomic properties by examining the relative placement of elements on the periodic table. - Describe the function of a designed material based upon its molecular properties through investigation. - support an argument for the conservation of mass in a chemical reaction. - explain changes in the composition of the nucleus of the atom and the release of energy during fission, fusion, or radioactive decay. - explain a design of a chemical system by specifying a change in conditions that would alter the amount of products at equilibrium. Collect, analyze and/or evaluate evidence - of chemical properties to explain the outcome of a chemical reaction. - to explain changes to chemical reaction rates. - of an explanation for products of a simple chemical reaction. -

PHYSICAL SCIENCE RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Motion Stability: Forces and Interactions	Identify or recognize - Newton's second law to describe force and motion relationships. - the concept of the conservation of momentum. - forces that act at a distance. - Newton's Law of Gravitation or recognize gravitational forces between objects. - producing a magnetic field or changing the magnetic field can produce electric current. - evidence of an electric current.	Identify a model to describe - or predict forces that act at a distance. - Newton's Law of Gravitation or predict gravitational forces between objects. - evidence of an electric current producing a magnetic field or changing the magnetic field can produce electric current. - force and motion relationships using Newton's second law. - the concept of the conservation of momentum.	Use evidence from a model to - compare the effects of forces on an object's motion. - support the claim there is conservation of momentum in a system. - describe and predict forces that act at a distance. - explain Newton's Law of Gravitation and predict gravitational forces between objects. - investigate an electric current producing a magnetic field or changing the magnetic field can produce electric current.	Analyze or evaluate evidence - to support Newton's second law of motion. - to explain the conservation of momentum. - to describe and predict forces that act a distance. - to explain Newton's Law of Gravitation and predict gravitational forces between objects. - from an investigation to explain an electric current producing a magnetic field or changing the magnetic field can produce electric current.
Energy	Identify or recognize - a change in energy of one component of the system. - how thermal energy is distributed in a closed system. - two objects interacting through electric or magnetic fields. - that energy can be a combination of motion and relative position of particles. - a design that involves the conversion of energy.	Describe or predict - a change in energy of one component of the system. - that energy can be a combination of motion and relative position of particles. - a design that involves the conversion of energy. - how thermal energy is distributed in a closed system. - the interaction between two objects through electric or magnetic fields.	Use models to - calculate a change in energy of one component of the system. - explain that energy can be a combination of motion and relative position of particles. - investigate how thermal energy is distributed in a closed system. - describe the interaction between two objects through electric or magnetic fields and the changes in energy due to the interaction. - refine a design that involves the conversion of energy.	Analyze or evaluate models to - calculate a change in energy of one component of the system. - explain that energy can be a combination of motion and relative position of particles. - make multiple refinements to a design that involves multiple conversions of energy. - to explain the interaction between two objects through electric or magnetic fields and the changes in energy due to the interaction. - investigate and explain how thermal energy is distributed in a closed system.

PHYSICAL SCIENCE RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Waves and Their Applications in Technologies for Information Transfer	Identify - how waves behave in different media. - information about electromagnetic radiation interacting with matter. - or describe claims regarding the effects different frequencies of electromagnetic radiation have when absorbed by matter. Recognize - wave and particle models of electromagnetic radiation.	Describe - how waves behave in different media. - wave and particle models of electromagnetic radiation. - information about electromagnetic radiation interacting with matter. Explain claims - regarding the effects different frequencies of electromagnetic radiation have when absorbed by matter.	Use models to - explain how waves behave in different media. - describe light through wave particles. - electromagnetic radiation interacting with matter. Analyze or evaluate claims - regarding the effects different frequencies of electromagnetic radiation have when absorbed by matter.	Develop, analyze, or evaluate models to - explain how waves behave in different media. - predict the behavior of light. - explain electromagnetic radiation interacting with matter. Develop claims - regarding the effects different frequencies of electromagnetic radiation have when absorbed by matter.

PHYSICAL SCIENCE RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Earth and Space Science	**Below Basic** A student who has reached the level of *Below Basic* level is to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Earth's Place in the Universe	Identify - the relationship between star properties and released energy. - the big bang theory. - the present orbital motions of objects in the solar system. - tectonic-plate movements that describe the relative ages of different materials on Earth. - data used to describe Earth's formation or early history. - physical processes on Earth's surface and within earth that shape Earth's features over time and space.	Describe - the relationship between star properties and released energy. - the present orbital motions of objects in the solar system. - how physical processes on Earth's surface and within earth shape Earth's features over time and space. - data used to summarize Earth's formation or early history. - evidence to summarize the big bang theory. - tectonic-plate movements to predict the relative ages of different materials on Earth.	Use a model to - explain the relationship between star properties and released energy. - support the big bang theory. - predict orbital motions of objects in the solar system. - describe evidence of tectonic-plate movements to predict the ages of different materials on Earth. - explain Earth's formation or early history. - explain how physical processes on Earth's surface and within earth shape Earth's features over time and space.	Analyze and/or evaluate models to - explain the relationship between star properties and released energy. - explain predictions of orbital motions of objects in the solar system. - synthesize astronomical evidence to support the big bang theory. - explain how physical processes on Earth's surface and within earth shape Earth's features over time and space. - support an argument about Earth's formation or early history. - evaluate evidence for the ages of different materials on Earth through tectonic-plate movements.

PHYSICAL SCIENCE RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Engineering and Technology Science	**Below Basic** A student who has reached the level of *Below Basic* level is to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Engineering Design	Recognize - criteria or constraints for solutions to a major global challenge. - an engineering problem. Identify - the needs or trade-offs of an engineering design. - a solution to a design problem.	Specify/describe - criteria and constraints for solutions to a major global challenge that account for societal needs and wants. Identify - that engineering problems can be broken down into smaller problems. - the needs and trade-offs of an engineering design. - the most appropriate solution to a design problem.	Use evidence to describe - criteria and constraints for solutions to a major global challenge that account for societal needs and wants. Use models to - design solutions to smaller problems in the context of a larger problem. - prioritize needs and trade-offs of an engineering design to evaluate a complex, real-world problem. - explain the most appropriate solution to a design problem.	Analyze or evaluate evidence - of criteria and constraints for solutions to a major global challenge that account for societal needs and wants. - for solutions to smaller problems in the context of a larger problem. - to argue for the most appropriate solution to a design problem. Communicate with evidence - prioritized needs and trade-offs of an engineering design to optimize a solution to a complex, real-world problem.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

ADVANCED

A student performing at Advanced effectively, consistently, and appropriately applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses information to evaluate patterns in data and revise models that support scientific claims, explain relationships among variables, and predict, based on scientific principles and reasoning, how the variables will change over time. The student revises the design of investigations in order to collect data that can describe quantitative relationships among variables. The student analyzes patterns in data to determine which solution best meets the criteria and constraints of a problem. The student uses data, mathematical and computational thinking, and scientific principles to construct explanations of scientific processes and arguments about stability and change within systems.

PROFICIENT

A student performing at Proficient effectively applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student develops models and uses information and patterns in data to support scientific arguments, describe relationships among variables, and predict how the variables will change over time. The student plans investigations to determine proportional relationships among variables. The student analyzes patterns in data to evaluate how well a solution meets the criteria and constraints of a problem. The student uses data, mathematical and computational thinking, and scientific principles to construct explanations of scientific processes and arguments about how systems and system parts will change over time.

BASIC

A student performing at Basic applies, with support, science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student uses models, information, and patterns in data to support scientific arguments, identify the relationship between two variables, and make predictions about how changes to one variable will affect other variables. The student describes the data to collect in an investigation in order to identify proportional relationships among variables. The student uses patterns in data to identify a solution that meets given criteria and constraints of a problem. The student uses data, basic algebraic thinking, and scientific principles to support explanations of scientific processes and arguments about how systems and system parts will change over time.

BELOW BASIC

A student performing at Below Basic seldom applies science and engineering practices to explain phenomena and design solutions to problems in the natural and the designed world. The student occasionally identifies models, information, and patterns in data to describe relationships between two variables and make predictions about how changes to one variable will affect other variables. The student infrequently recognizes trends in the data collected during an investigation in order to identify the relationships among variables. The student can sometimes use patterns in data to identify a solution to a problem. The student occasionally uses data and basic algebraic thinking to explain how systems and system parts change over time.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Life Science	**Below Basic** A student who has reached the level of *Below Basic* is able to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
From Molecules to Organisms: Structure and Processes	Recognize - that DNA controls specialized cell functions. - the process of mitosis, cellular division, or differentiation. - a model of photosynthesis that transforms light energy into chemical energy. - a model of cellular respiration breaks down molecules resulting in a net transfer of energy.	Identify evidence of or a model of - how DNA sequences relate to specialized cell functions. - the interaction of cells or organs provide specific functions. - that describes the process of mitosis, cellular division, or differentiation in producing and maintaining complex organisms. - that describes photosynthesis transforms light energy into chemical energy. - that describes cellular respiration breaks down molecules resulting in a net transfer of energy.	Develop or use models to explain - how DNA sequences relate to specialized cell functions. - how interacting cells or organs provide specific functions. - photosynthesis transforms light energy into chemical energy. - cellular respiration breaks down molecules resulting in a net transfer of energy. - the process of mitosis, cellular division, or differentiation in producing and maintaining complex organisms.	Analyze or evaluate evidence of a model to explain - how DNA sequences relate to specialized cell functions. - to explain how interacting cells or organs provide specific functions. - photosynthesis transforms light energy into chemical energy. - cellular respiration breaks down molecules resulting in a net transfer of energy. - the process of mitosis, cellular division, or differentiation in producing and maintaining complex organisms.
	Identify - how interacting cells or organs provide specific functions. - or recognize an explanation of macromolecule composition.	Use data to identify - how life functions rely upon homeostasis.	Investigate - how life functions rely upon feedback mechanisms in homeostasis.	Investigate and evaluate evidence of - how life functions rely upon feedback mechanisms in homeostasis.
	Observe - life functions that rely upon homeostasis.	Revise an explanation of - macromolecule composition.	Create an explanation of - macromolecule composition.	Evaluate an explanation and provide evidence of - macromolecule composition.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Ecosystems: Interactions, Energy, and Dynamics	Describe - the transformation in plants of light into chemical energy. - how matter and energy found in food molecules are used in organisms. - the biological processes that cycle carbon and energy within Earth systems. - physical or biological changes that affect ecosystem conditions and stability.	Identify a model to describe - the transformation in plants of light into chemical energy. - how matter and energy found in food molecules are used in organisms. - the biological processes that cycle carbon and energy within Earth systems - physical or biological changes that affect ecosystem conditions and stability.	Use models to explain - the transformation in plants of light into chemical energy. - how matter and energy found in food molecules are used in organisms. - biological processes that cycle carbon and energy within Earth systems. - factors affecting biodiversity and ecosystem populations.	Analyze or evaluate evidence of models to explain - the transformation in plants of light into chemical energy. - how matter and energy found in food molecules are used in organisms. - the biological processes that cycle carbon and energy within Earth systems. - to explain factors affecting biodiversity and ecosystem populations.
	Explain - factors affecting biodiversity of ecosystem populations. - how to minimize human impacts on the environment and biodiversity.	Identify evidence to describe - factors affecting biodiversity of ecosystem populations. - a design that minimizes human impacts on the environment and biodiversity.	Evaluate - evidence of complex physical or biological changes that affect ecosystem conditions and stability. - designs that minimize human impacts on the environment and biodiversity.	Gather, analyze and/or evaluate evidence to communicate - complex physical or biological changes that affect ecosystem conditions and stability. - designs that minimize human impacts on the environment and biodiversity.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Heredity: Inheritance and Variation of Traits	Recognize - differences in the complexity of organisms caused by cellular divisions. - DNA data to describe genetic variation in individuals of a population. - the probability of a trait in a population. - similarities and differences of asexual and sexual reproduction.	Identify models to describe - differences in the complexity of organisms caused by cellular divisions. - variation in organisms as a result of asexual or sexual reproduction. - DNA findings explain genetic variation in individuals of a population. - the probability of a trait in a population.	Use a model to explain - differences in the complexity of organisms caused by cellular divisions. - variation in organisms as a result of asexual or sexual reproduction. - the variation and distribution of traits in a population. - the cause of genetic variation in individuals and in populations.	Construct and/or evaluate a model to explain or predict - differences in the complexity of organisms caused by cellular divisions. - variation in organisms as a result of asexual or sexual reproduction. - variation and distribution of traits in a population as they relate to genetic and environmental factors. - for the cause of genetic variation in individuals and in populations.
	Identify a claim about - the causes of inheritable genetic variation.	Make a claim about - the causes of inheritable genetic variation.	Use evidence to make and defend a claim about - the causes of inheritable genetic variation.	Evaluate evidence to make and defend a claim - about the causes of inheritable genetic variation.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Biological Evolution: Unity and Diversity	Identify or describe - common ancestry and biological evolution. - similarities in embryological development across multiple species. - the likelihood of an organism with advantageous heritable characteristic will increase. - environmental changes that affect species populations over time. - a solution to negative effects of human impact on biodiversity.	Identify evidence to describe - common ancestry and biological evolution. - similarities in embryological development across multiple species. - ecological and genetic factors related to evolutionary processes. - the adaptation of populations through natural selection. - the likelihood of an organism with advantageous heritable characteristic will increase. - environmental changes that affect species populations over time.	Use evidence to communicate - common ancestry and biological evolution. - similarities in embryological development across multiple species. - that ecological and genetic factors result in evolutionary processes. - the likelihood of an organism with advantageous heritable characteristic will increase. - adaptations of populations through natural selection. - or argue that environmental changes affect species populations over time.	Collect, analyze and/or evaluate evidence to - determine which best explains common ancestry and biological evolution. - explain relationships in similarities evident in embryological development across multiple spaces. - explain ecological and genetic factors result in evolutionary processes. - explain the likelihood of an organism with advantageous heritable characteristic will increase. - support the adaptation of populations through natural selection. - explain that environmental changes affect species populations over time.
	Explain - ecological and genetic factors related to evolutionary processes. - the adaptation of populations through natural selection.	Use a model to describe - a solution to negative effects of human impact on biodiversity.	Develop a model to explain - a solution to negative effects of human impact on biodiversity.	Revise or evaluate a model to - explain a solution to negative effects of human impact on biodiversity.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Earth and Space Science	**Below Basic** A student who has reached the level of *Below Basic* is to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Earth and Human Activity	Describe - the effects of natural resources or natural hazards on human activity. - how human activity affects Earth Systems	Identify evidence within a model to describe - the effects of natural resources or natural hazards on human activity. - how human activity affects Earth Systems	Use a model to - explain the effects of natural resources or natural hazards on human activity. - predict how human activity affects Earth Systems	Develop, analyze or evaluate a model from evidence to - explain the effects of natural resources or natural hazards on human activity. - predict how human activity affects Earth Systems
	Identify - an example of coevolution of Earth's systems and life on Earth. - a solution that reduces human impacts on natural systems.	Identify an argument based on evidence - to describe coevolution of Earth's systems and life on Earth. -	Construct an argument based on evidence - to describe coevolution of Earth's systems and life on Earth.	Analyze or evaluate an argument - to explain coevolution of Earth's systems and life on Earth. -
	Recognize - the impacts of human use of natural resources.	Describe evidence of - the impacts of human use of natural resources. - a solution that reduces human impacts on natural systems.	Evaluate - the cost-benefit design solutions for the use of natural resources. - a solution that reduces human impacts on natural systems.	Evaluate - and optimize the cost-benefit design solutions for the use of natural resources. - refine and communicate solutions that reduces human impacts on natural systems.

BIOLOGY RANGE PERFORMANCE LEVEL DESCRIPTORS - SCIENCE

Engineering and Technology Science	**Below Basic** A student who has reached the level of *Below Basic* is to successfully address some, but not all, of the following:	**Basic** A student who has reached the level of *Basic* is able to successfully address some, but not all, of the following:	**Proficient** A student who has reached the level of *Proficient* is able to successfully address some, but not all, of the following:	**Advanced** A student who has reached the level of *Advanced* is able to successfully address some, but not all, of the following:
Engineering Design	Recognize - criteria or constraints for solutions to a major global challenge. - an engineering problem. - the needs or trade-offs of an engineering design.	Specify evidence of - criteria and constraints for solutions to a major global challenge that account for societal needs and wants. - engineering problems can be broken down into smaller problems. - the needs and trade-offs of an engineering design.	Use evidence to describe - criteria and constraints for solutions to a major global challenge that account for societal needs and wants. - design solutions to smaller problems in the context of a larger problem. - prioritize needs and trade-offs of an engineering design to evaluate a complex, real-world problem. -	Analyze or evaluate evidence - of criteria and constraints for solutions to a major global challenge that account for societal needs and wants. - for solutions to smaller problems in the context of a larger problem. - prioritized needs and trade-offs of an engineering design to optimize a solution to a complex, real-world problem.
	Identify - a solution to a design problem.	Identify a model for - the most appropriate solution to a design problem.	Use models to - explain the most appropriate solution to a design problem.	Communicate with evidence - to argue for the most appropriate solution to a design problem.

F

Detailed Reports of Participants' Judgments

Missouri Grade 5 Science
Round 1 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	401	10	50	67
1	402	15	43	60
1	403	11	49	60
1	404	26	45	57
1	405	28	45	58
1	406	21	41	62
1	407	10	43	55
2	408	13	47	55
2	409	17	41	58
2	410	26	47	60
2	411	20	41	60
2	412	19	46	61
2	413	22	39	52

Overall	Median	19	45	60
	25th %ile	12	41	56
	75th %ile	24	47	60.5
	Minimum	10	39	52
	Maximum	28	50	67

Missouri Grade 5 Science
Round 1 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	401	247	317	374
1	402	263	308	344
1	403	255	316	344
1	404	285	311	337
1	405	291	311	339
1	406	279	306	358
1	407	247	308	327
2	408	261	312	327
2	409	268	306	339
2	410	285	312	344
2	411	277	306	344
2	412	275	311	352
2	413	281	303	319

Overall	Median	275	311	344
	25th %ile	258	306	332
	75th %ile	283	312	348
	Minimum	247	303	319
	Maximum	291	317	374

Missouri Grade 5 Science
Round 1 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	15	45	60
Median	2	19.5	43.5	59
Median	Overall	19	45	60
25th %ile	1	10	43	57
25th %ile	2	16	40.5	54.25
25th %ile	Overall	12	41	56
75th %ile	1	26	49	62
75th %ile	2	23	47	60.25
75th %ile	Overall	24	47	60.5
Minimum	1	10	41	55
Minimum	2	13	39	52
Minimum	Overall	10	39	52
Maximum	1	28	50	67
Maximum	2	26	47	61
Maximum	Overall	28	50	67

Overall	Median	19	45	60
	25th %ile	12	41	56
	75th %ile	24	47	60.5
	Minimum	10	39	52
	Maximum	28	50	67

Missouri Grade 5 Science
Round 1 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	263	311	344
Median	2	276	309	341
Median	Overall	275	311	344
25th %ile	1	247	308	337
25th %ile	2	266	305	325
25th %ile	Overall	258	306	332
75th %ile	1	285	316	358
75th %ile	2	282	312	346
75th %ile	Overall	283	312	348
Minimum	1	247	306	327
Minimum	2	261	303	319
Minimum	Overall	247	303	319
Maximum	1	291	317	374
Maximum	2	285	312	352
Maximum	Overall	291	317	374

Overall	Median	275	311	344
	25th %ile	258	306	332
	75th %ile	283	312	348
	Minimum	247	303	319
	Maximum	291	317	374

Missouri Grade 5 Science
Round 1 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	15	45	60
2	19.5	43.5	59
Overall	19	45	60

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	26.0	32.1	28.3	13.5

Missouri Grade 5 Science
Round 2 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	401	17	44	55
1	402	18	43	60
1	403	19	44	60
1	404	19	43	57
1	405	19	45	60
1	406	19	41	58
1	407	15	45	59
2	408	17	47	61
2	409	17	46	58
2	410	19	48	61
2	411	17	35	58
2	412	19	47	61
2	413	17	46	55

Overall	Median	18	45	59
	25th %ile	17	43	57.5
	75th %ile	19	46.5	60.5
	Minimum	15	35	55
	Maximum	19	48	61

Missouri Grade 5 Science
Round 2 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	401	268	310	327
1	402	272	308	344
1	403	275	310	344
1	404	275	308	337
1	405	275	311	344
1	406	275	306	339
1	407	263	311	341
2	408	268	312	352
2	409	268	311	339
2	410	275	314	352
2	411	268	300	339
2	412	275	312	352
2	413	268	311	327

Overall	Median	272	311	341
	25th %ile	268	308	338
	75th %ile	275	312	348
	Minimum	263	300	327
	Maximum	275	314	352

Missouri Grade 5 Science
Round 2 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	19	44	59
Median	2	17	46.5	59.5
Median	Overall	18	45	59
25th %ile	1	17	43	57
25th %ile	2	17	43.25	57.25
25th %ile	Overall	17	43	57.5
75th %ile	1	19	45	60
75th %ile	2	19	47.25	61
75th %ile	Overall	19	46.5	60.5
Minimum	1	15	41	55
Minimum	2	17	35	55
Minimum	Overall	15	35	55
Maximum	1	19	45	60
Maximum	2	19	48	61
Maximum	Overall	19	48	61

Overall	Median	18	45	59
	25th %ile	17	43	57.5
	75th %ile	19	46.5	60.5
	Minimum	15	35	55
	Maximum	19	48	61

Missouri Grade 5 Science
Round 2 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	275	310	341
Median	2	268	312	346
Median	Overall	272	311	341
25th %ile	1	268	308	337
25th %ile	2	268	309	336
25th %ile	Overall	268	308	338
75th %ile	1	275	311	344
75th %ile	2	275	312	352
75th %ile	Overall	275	312	348
Minimum	1	263	306	327
Minimum	2	268	300	327
Minimum	Overall	263	300	327
Maximum	1	275	311	344
Maximum	2	275	314	352
Maximum	Overall	275	314	352

Overall	Median	272	311	341
	25th %ile	268	308	338
	75th %ile	275	312	348
	Minimum	263	300	327
	Maximum	275	314	352

Missouri Grade 5 Science
Round 2 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	19	44	59
2	17	46.5	59.5
Overall	18	45	59

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	24.0	34.2	26.3	15.6

Missouri Grade 5 Science
Round 3 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	401	19	44	60
1	402	18	43	60
1	403	19	44	60
1	404	19	43	58
1	405	19	43	60
1	406	19	41	58
1	407	19	45	59
2	408	17	47	61
2	409	19	47	60
2	410	19	48	61
2	411	18	42	59
2	412	19	47	61
2	413	17	47	58

Overall	Median	19	44	60
	25th %ile	18	43	58.5
	75th %ile	19	47	60.5
	Minimum	17	41	58
	Maximum	19	48	61

Missouri Grade 5 Science
Round 3 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	401	275	310	344
1	402	272	308	344
1	403	275	310	344
1	404	275	308	339
1	405	275	308	344
1	406	275	306	339
1	407	275	311	341
2	408	268	312	352
2	409	275	312	344
2	410	275	314	352
2	411	272	306	341
2	412	275	312	352
2	413	268	312	339

Overall	Median	275	310	344
	25th %ile	272	308	340
	75th %ile	275	312	348
	Minimum	268	306	339
	Maximum	275	314	352

Missouri Grade 5 Science
Round 3 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	19	43	60
Median	2	18.5	47	60.5
Median	Overall	19	44	60
25th %ile	1	19	43	58
25th %ile	2	17	45.75	58.75
25th %ile	Overall	18	43	58.5
75th %ile	1	19	44	60
75th %ile	2	19	47.25	61
75th %ile	Overall	19	47	60.5
Minimum	1	18	41	58
Minimum	2	17	42	58
Minimum	Overall	17	41	58
Maximum	1	19	45	60
Maximum	2	19	48	61
Maximum	Overall	19	48	61

Overall	Median	19	44	60
	25th %ile	18	43	58.5
	75th %ile	19	47	60.5
	Minimum	17	41	58
	Maximum	19	48	61

Missouri Grade 5 Science
Round 3 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	275	308	344
Median	2	273	312	348
Median	Overall	275	310	344
25th %ile	1	275	308	339
25th %ile	2	268	310	341
25th %ile	Overall	272	308	340
75th %ile	1	275	310	344
75th %ile	2	275	312	352
75th %ile	Overall	275	312	348
Minimum	1	272	306	339
Minimum	2	268	306	339
Minimum	Overall	268	306	339
Maximum	1	275	311	344
Maximum	2	275	314	352
Maximum	Overall	275	314	352

Overall	Median	275	310	344
	25th %ile	272	308	340
	75th %ile	275	312	348
	Minimum	268	306	339
	Maximum	275	314	352

Missouri Grade 5 Science
Round 3 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	19	43	60
2	18.5	47	60.5
Overall	19	44	60

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	26.0	31.1	29.3	13.5

Missouri Grade 8 Science
Round 1 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	421	16	32	51
1	422	20	39	59
1	423	18	35	54
1	424	18	39	57
1	425	20	41	60
1	426	23	42	55
2	427	22	42	49
2	428	22	41	53
2	429	22	39	60
2	430	15	36	47
2	431	22	48	59

Overall	Median	20	39	55
	25th %ile	18	36	51
	75th %ile	22	42	59
	Minimum	15	32	47
	Maximum	23	48	60

Missouri Grade 8 Science
Round 1 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	421	462	497	518
1	422	476	504	540
1	423	469	500	526
1	424	469	504	534
1	425	476	506	544
1	426	481	508	528
2	427	480	508	516
2	428	480	506	523
2	429	480	504	544
2	430	461	501	515
2	431	480	515	540

Overall	Median	476	504	528
	25th %ile	469	501	518
	75th %ile	480	508	540
	Minimum	461	497	515
	Maximum	481	515	544

Missouri Grade 8 Science
Round 1 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	19	39	56
Median	2	22	41	53
Median	Overall	20	39	55
25th %ile	1	17.5	34.25	53.25
25th %ile	2	18.5	37.5	48
25th %ile	Overall	18	36	51
75th %ile	1	20.75	41.25	59.25
75th %ile	2	22	45	59.5
75th %ile	Overall	22	42	59
Minimum	1	16	32	51
Minimum	2	15	36	47
Minimum	Overall	15	32	47
Maximum	1	23	42	60
Maximum	2	22	48	60
Maximum	Overall	23	48	60

Overall	Median	20	39	55
	25th %ile	18	36	51
	75th %ile	22	42	59
	Minimum	15	32	47
	Maximum	23	48	60

Missouri Grade 8 Science
Round 1 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	472	504	531
Median	2	480	506	523
Median	Overall	476	504	528
25th %ile	1	467	499	524
25th %ile	2	471	502	515
25th %ile	Overall	469	501	518
75th %ile	1	477	506	541
75th %ile	2	480	512	542
75th %ile	Overall	480	508	540
Minimum	1	462	497	518
Minimum	2	461	501	515
Minimum	Overall	461	497	515
Maximum	1	481	508	544
Maximum	2	480	515	544
Maximum	Overall	481	515	544

Overall	Median	476	504	528
	25th %ile	469	501	518
	75th %ile	480	508	540
	Minimum	461	497	515
	Maximum	481	515	544

Missouri Grade 8 Science
Round 1 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	19	39	56
2	22	41	53
Overall	20	39	55

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	25.6	24.5	24.2	25.8

Missouri Grade 8 Science
Round 2 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	421	18	34	54
1	422	20	39	59
1	423	18	35	56
1	424	20	39	58
1	425	20	39	57
1	426	21	39	59
2	427	21	39	53
2	428	20	39	58
2	429	20	37	58
2	430	20	40	58
2	431	22	41	59

Overall	Median	20	39	58
	25th %ile	20	37	56
	75th %ile	21	39	59
	Minimum	18	34	53
	Maximum	22	41	59

Missouri Grade 8 Science
Round 2 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	421	469	498	526
1	422	476	504	540
1	423	469	500	530
1	424	476	504	537
1	425	476	504	534
1	426	479	504	540
2	427	479	504	523
2	428	476	504	537
2	429	476	501	537
2	430	476	505	537
2	431	480	506	540

Overall	Median	476	504	537
	25th %ile	476	501	530
	75th %ile	479	504	540
	Minimum	469	498	523
	Maximum	480	506	540

Missouri Grade 8 Science
Round 2 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	20	39	57.5
Median	2	20	39	58
Median	Overall	20	39	58
25th %ile	1	18	34.75	55.5
25th %ile	2	20	38	55.5
25th %ile	Overall	20	37	56
75th %ile	1	20.25	39	59
75th %ile	2	21.5	40.5	58.5
75th %ile	Overall	21	39	59
Minimum	1	18	34	54
Minimum	2	20	37	53
Minimum	Overall	18	34	53
Maximum	1	21	39	59
Maximum	2	22	41	59
Maximum	Overall	22	41	59

Overall	Median	20	39	58
	25th %ile	20	37	56
	75th %ile	21	39	59
	Minimum	18	34	53
	Maximum	22	41	59

Missouri Grade 8 Science
Round 2 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	476	504	536
Median	2	476	504	537
Median	Overall	476	504	537
25th %ile	1	469	500	529
25th %ile	2	476	503	530
25th %ile	Overall	476	501	530
75th %ile	1	477	504	540
75th %ile	2	479	505	538
75th %ile	Overall	479	504	540
Minimum	1	469	498	526
Minimum	2	476	501	523
Minimum	Overall	469	498	523
Maximum	1	479	504	540
Maximum	2	480	506	540
Maximum	Overall	480	506	540

Overall	Median	476	504	537
	25th %ile	476	501	530
	75th %ile	479	504	540
	Minimum	469	498	523
	Maximum	480	506	540

Missouri Grade 8 Science
Round 2 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	20	39	57.5
2	20	39	58
Overall	20	39	58

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	25.6	24.5	31.9	18.1

Missouri Grade 8 Science
Round 3 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	421	16	36	57
1	422	15	37	59
1	423	16	35	56
1	424	15	38	58
1	425	15	37	57
1	426	15	37	55
2	427	15	37	58
2	428	15	37	58
2	429	16	37	58
2	430	15	37	58
2	431	13	37	59

Overall	Median	15	37	58
	25th %ile	15	37	57
	75th %ile	16	37	58
	Minimum	13	35	55
	Maximum	16	38	59

Missouri Grade 8 Science
Round 3 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	421	462	501	534
1	422	461	501	540
1	423	462	500	530
1	424	461	503	537
1	425	461	501	534
1	426	461	501	528
2	427	461	501	537
2	428	461	501	537
2	429	462	501	537
2	430	461	501	537
2	431	456	501	540

Overall	Median	461	501	537
	25th %ile	461	501	534
	75th %ile	462	501	537
	Minimum	456	500	528
	Maximum	462	503	540

Missouri Grade 8 Science
Round 3 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	15	37	57
Median	2	15	37	58
Median	Overall	15	37	58
25th %ile	1	15	35.75	55.75
25th %ile	2	14	37	58
25th %ile	Overall	15	37	57
75th %ile	1	16	37.25	58.25
75th %ile	2	15.5	37	58.5
75th %ile	Overall	16	37	58
Minimum	1	15	35	55
Minimum	2	13	37	58
Minimum	Overall	13	35	55
Maximum	1	16	38	59
Maximum	2	16	37	59
Maximum	Overall	16	38	59

Overall	Median	15	37	58
	25th %ile	15	37	57
	75th %ile	16	37	58
	Minimum	13	35	55
	Maximum	16	38	59

Missouri Grade 8 Science
Round 3 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	461	501	534
Median	2	461	501	537
Median	Overall	461	501	537
25th %ile	1	461	501	529
25th %ile	2	459	501	537
25th %ile	Overall	461	501	534
75th %ile	1	462	502	538
75th %ile	2	462	501	538
75th %ile	Overall	462	501	537
Minimum	1	461	500	528
Minimum	2	456	501	537
Minimum	Overall	456	500	528
Maximum	1	462	503	540
Maximum	2	462	501	540
Maximum	Overall	462	503	540

Overall	Median	461	501	537
	25th %ile	461	501	534
	75th %ile	462	501	537
	Minimum	456	500	528
	Maximum	462	503	540

Missouri Grade 8 Science
Round 3 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	15	37	57
2	15	37	58
Overall	15	37	58

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	16.5	30.4	35.0	18.1

Missouri Physical Science
Round 1 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	446	19	51	68
1	447	23	50	73
1	448	15	56	77
1	449	12	49	64
1	460	21	57	73
2	441	18	57	71
2	442	17	59	71
2	443	19	57	78
2	444	13	17	38
2	445	17	56	74
2	451	20	57	77

Overall		Basic	Proficient	Advanced
	Median	18.0	56.0	73.0
	25th %ile	16.0	50.5	69.5
	75th %ile	19.5	57.0	75.5
	Minimum	12.0	17.0	38.0
	Maximum	23.0	59.0	78.0

Missouri Physical Science
Round 1 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	446	383	398	408
1	447	385	398	415
1	448	381	400	425
1	449	378	397	404
1	460	384	400	415
2	441	383	400	414
2	442	382	401	414
2	443	383	400	431
2	444	380	382	391
2	445	382	400	417
2	451	383	400	425

Overall		Basic	Proficient	Advanced
	Median	383	400	415
	25th %ile	382	398	412
	75th %ile	383	400	421
	Minimum	378	382	391
	Maximum	385	401	431

Missouri Physical Science
Round 1 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	19	51	73
Median	2	17.5	57	72.5
Median	Overall	18	56	73
25th %ile	1	15	50	68
25th %ile	2	17	56.25	71
25th %ile	Overall	16	50.5	69.5
75th %ile	1	21	56	73
75th %ile	2	18.75	57	76.25
75th %ile	Overall	19.5	57	75.5
Minimum	1	12	49	64
Minimum	2	13	17	38
Minimum	Overall	12	17	38
Maximum	1	23	57	77
Maximum	2	20	59	78
Maximum	Overall	23	59	78

Overall		Basic	Proficient	Advanced
	Median	18.0	56.0	73.0
	25th %ile	16.0	50.5	69.5
	75th %ile	19.5	57.0	75.5
	Minimum	12.0	17.0	38.0
	Maximum	23.0	59.0	78.0

Missouri Physical Science
Round 1 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	383	398	415
Median	2	383	400	415
Median	Overall	383	400	415
25th %ile	1	381	398	408
25th %ile	2	382	400	414
25th %ile	Overall	382	398	412
75th %ile	1	384	400	415
75th %ile	2	383	400	425
75th %ile	Overall	383	400	421
Minimum	1	378	397	404
Minimum	2	380	382	391
Minimum	Overall	378	382	391
Maximum	1	385	400	425
Maximum	2	383	401	431
Maximum	Overall	385	401	431

Overall		Basic	Proficient	Advanced
	Median	383	400	415
	25th %ile	382	398	412
	75th %ile	383	400	421
	Minimum	378	382	391
	Maximum	385	401	431

Missouri Physical Science
Round 1 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	19	51	73
2	17.5	57	72.5
Overall	18	56	73

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	19.2%	42.9%	28.2%	9.7%

Missouri Physical Science
Round 2 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	446	12	51	69
1	447	12	51	69
1	448	12	51	69
1	449	12	51	69
1	460	17	57	73
2	441	17	57	74
2	442	17	56	75
2	443	17	49	74
2	444	16	57	69
2	445	17	56	74
2	451	17	57	74

Overall		Basic	Proficient	Advanced
	Median	17.0	56.0	73.0
	25th %ile	12.0	51.0	69.0
	75th %ile	17.0	57.0	74.0
	Minimum	12.0	49.0	69.0
	Maximum	17.0	57.0	75.0

Missouri Physical Science
Round 2 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	446	378	398	410
1	447	378	398	410
1	448	378	398	410
1	449	378	398	410
1	460	382	400	415
2	441	382	400	417
2	442	382	400	419
2	443	382	397	417
2	444	382	400	410
2	445	382	400	417
2	451	382	400	417
Overall		**Basic**	**Proficient**	**Advanced**
	Median	382	400	415
	25th %ile	378	398	410
	75th %ile	382	400	417
	Minimum	378	397	410
	Maximum	382	400	419

Missouri Physical Science
Round 2 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	12	51	69
Median	2	17	56.5	74
Median	Overall	17	56	73
25th %ile	1	12	51	69
25th %ile	2	17	56	74
25th %ile	Overall	12	51	69
75th %ile	1	12	51	69
75th %ile	2	17	57	74
75th %ile	Overall	17	57	74
Minimum	1	12	51	69
Minimum	2	16	49	69
Minimum	Overall	12	49	69
Maximum	1	17	57	73
Maximum	2	17	57	75
Maximum	Overall	17	57	75

Overall		Basic	Proficient	Advanced
	Median	17.0	56.0	73.0
	25th %ile	12.0	51.0	69.0
	75th %ile	17.0	57.0	74.0
	Minimum	12.0	49.0	69.0
	Maximum	17.0	57.0	75.0

Missouri Physical Science
Round 2 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	378	398	410
Median	2	382	400	417
Median	Overall	382	400	415
25th %ile	1	378	398	410
25th %ile	2	382	400	417
25th %ile	Overall	382	398	410
75th %ile	1	378	398	410
75th %ile	2	382	400	417
75th %ile	Overall	382	400	417
Minimum	1	378	398	410
Minimum	2	382	397	410
Minimum	Overall	378	397	410
Maximum	1	382	400	415
Maximum	2	382	400	419
Maximum	Overall	382	400	419

Overall		Basic	Proficient	Advanced
	Median	382	400	415
	25th %ile	378	398	410
	75th %ile	382	400	417
	Minimum	378	397	410
	Maximum	382	400	419

Missouri Physical Science
Round 2 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	12	51	69
2	17	56.5	74
Overall	17	56	73

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	19.2%	42.9%	28.2%	9.7%

Missouri Physical Science
Round 3 Bookmark Placements

Table	Participant	Basic	Proficient	Advanced
1	446	11	49	69
1	447	11	49	69
1	448	11	49	69
1	449	11	49	69
1	460	11	51	71
2	441	17	57	74
2	442	17	56	74
2	443	17	57	74
2	444	17	56	74
2	445	17	56	74
2	451	17	57	74

Overall		Basic	Proficient	Advanced
	Median	17.0	56.0	74.0
	25th %ile	11.0	49.0	69.0
	75th %ile	17.0	56.5	74.0
	Minimum	11.0	49.0	69.0
	Maximum	17.0	57.0	74.0

Missouri Physical Science
Round 3 Cut Scores

Table	Participant	Basic	Proficient	Advanced
1	446	375	397	410
1	447	375	397	410
1	448	375	397	410
1	449	375	397	410
1	460	375	398	414
2	441	382	400	417
2	442	382	400	417
2	443	382	400	417
2	444	382	400	417
2	445	382	400	417
2	451	382	400	417
Overall		**Basic**	**Proficient**	**Advanced**
	Median	382	400	417
	25th %ile	375	397	410
	75th %ile	382	400	417
	Minimum	375	397	410
	Maximum	382	400	417

Missouri Physical Science
Round 3 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	11	49	69
Median	2	17	56.5	74
Median	Overall	17	56	74
25th %ile	1	11	49	69
25th %ile	2	17	56	74
25th %ile	Overall	11	49	69
75th %ile	1	11	49	69
75th %ile	2	17	57	74
75th %ile	Overall	17	56.5	74
Minimum	1	11	49	69
Minimum	2	17	56	74
Minimum	Overall	11	49	69
Maximum	1	11	51	71
Maximum	2	17	57	74
Maximum	Overall	17	57	74

Overall		Basic	Proficient	Advanced
	Median	17.0	56.0	74.0
	25th %ile	11.0	49.0	69.0
	75th %ile	17.0	56.5	74.0
	Minimum	11.0	49.0	69.0
	Maximum	17.0	57.0	74.0

Missouri Physical Science
Round 3 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	375	397	410
Median	2	382	400	417
Median	Overall	382	400	417
25th %ile	1	375	397	410
25th %ile	2	382	400	417
25th %ile	Overall	375	397	410
75th %ile	1	375	397	410
75th %ile	2	382	400	417
75th %ile	Overall	382	400	417
Minimum	1	375	397	410
Minimum	2	382	400	417
Minimum	Overall	375	397	410
Maximum	1	375	398	414
Maximum	2	382	400	417
Maximum	Overall	382	400	417

Overall		Basic	Proficient	Advanced
	Median	382	400	417
	25th %ile	375	397	410
	75th %ile	382	400	417
	Minimum	375	397	410
	Maximum	382	400	417

Missouri Physical Science
Round 3 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	11	49	69
2	17	56.5	74
Overall	17	56	74

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	19.2%	42.9%	28.2%	9.7%

Missouri Biology
Round 1 Bookmark Placements

Table	Panelist	Basic	Proficient	Advanced
1	463	26	72	84
1	464	24	74	85
1	468	45	57	84
1	469	28	74	86
1	471	35	57	80
1	472	50	75	86
2	461	33	51	74
2	462	34	64	78
2	465	22	42	70
2	466	31	58	70
2	467	27	51	70
2	470	17	63	82
Overall		**Basic**	**Proficient**	**Advanced**
	Median	29.5	60.5	81
	25th %ile	25.5	55.5	73
	75th %ile	34.25	72.5	84.25
	Minimum	17	42	70
	Maximum	50	75	86

Missouri Biology
Round 1 Cut Scores

Table	Panelist	Basic	Proficient	Advanced
1	463	380	400	411
1	464	380	401	412
1	468	389	393	411
1	469	381	401	424
1	471	385	393	405
1	472	390	401	424
2	461	384	392	401
2	462	384	396	404
2	465	379	387	399
2	466	383	394	399
2	467	381	392	399
2	470	378	394	406
Overall		**Basic**	**Proficient**	**Advanced**
	Median	381	394	405
	25th %ile	380	392	401
	75th %ile	385	401	412
	Minimum	378	387	399
	Maximum	390	401	424

Missouri Biology
Round 1 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	31.5	73	84.5
Median	2	29	54.5	72
Median	Overall	29.5	60.5	81
25th %ile	1	26.5	60.75	84
25th %ile	2	23.25	51	70
25th %ile	Overall	25.5	55.5	73
75th %ile	1	42.5	74	85.75
75th %ile	2	32.5	61.75	77
75th %ile	Overall	34.25	72.5	84.25
Minimum	1	24	57	80
Minimum	2	17	42	70
Minimum	Overall	17	42	70
Maximum	1	50	75	86
Maximum	2	34	64	82
Maximum	Overall	50	75	86

Overall		Basic	Proficient	Advanced
	Median	29.5	60.5	81
	25th %ile	25.5	55.5	73
	75th %ile	34.25	72.5	84.25
	Minimum	17	42	70
	Maximum	50	75	86

Missouri Biology
Round 1 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	384	401	412
Median	2	381	392	400
Median	Overall	381	394	406
25th %ile	1	381	394	411
25th %ile	2	380	392	399
25th %ile	Overall	380	392	401
75th %ile	1	388	401	424
75th %ile	2	384	394	403
75th %ile	Overall	385	401	412
Minimum	1	380	393	405
Minimum	2	378	387	399
Minimum	Overall	378	387	399
Maximum	1	390	401	424
Maximum	2	384	396	406
Maximum	Overall	390	401	424

Overall		Basic	Proficient	Advanced
	Median	381	394	405
	25th %ile	380	392	401
	75th %ile	385	401	412
	Minimum	378	387	399
	Maximum	390	401	424

Missouri Biology
Round 1 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	31.5	73	84.5
2	29	54.5	72
Overall	29.5	60.5	81

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	19.1%	28.3%	27.3%	25.4%

Missouri Biology
Round 2 Bookmark Placements

Table	Panelist	Basic	Proficient	Advanced
1	463	23	70	80
1	464	24	70	84
1	468	28	64	84
1	469	28	72	84
1	471	29	72	80
1	472	26	75	86
2	461	23	51	70
2	462	34	64	81
2	465	25	51	71
2	466	25	63	82
2	467	27	58	73
2	470	22	63	82
Overall		**Basic**	**Proficient**	**Advanced**
	Median	25.5	64	81.5
	25th %ile	23.75	61.75	78.25
	75th %ile	28	70.5	84
	Minimum	22	51	70
	Maximum	34	75	86

Missouri Biology
Round 2 Cut Scores

Table	Panelist	Basic	Proficient	Advanced
1	463	380	399	405
1	464	380	399	411
1	468	381	396	411
1	469	381	400	411
1	471	381	400	405
1	472	380	401	424
2	461	380	392	399
2	462	384	396	406
2	465	380	392	400
2	466	380	394	406
2	467	381	394	401
2	470	379	394	406
Overall		**Basic**	**Proficient**	**Advanced**
	Median	380	396	406
	25th %ile	380	394	405
	75th %ile	381	400	411
	Minimum	379	392	399
	Maximum	384	401	424

Missouri Biology
Round 2 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	27	71	84
Median	2	25	60.5	77
Median	Overall	25.5	64	81.5
25th %ile	1	24.5	70	81
25th %ile	2	23.5	52.75	71.5
25th %ile	Overall	23.75	61.75	78.25
75th %ile	1	28	72	84
75th %ile	2	26.5	63	81.75
75th %ile	Overall	28	70.5	84
Minimum	1	23	64	80
Minimum	2	22	51	70
Minimum	Overall	22	51	70
Maximum	1	29	75	86
Maximum	2	34	64	82
Maximum	Overall	34	75	86

Overall		Basic	Proficient	Advanced
	Median	25.5	64	81.5
	25th %ile	23.75	61.75	78.25
	75th %ile	28	70.5	84
	Minimum	22	51	70
	Maximum	34	75	86

Missouri Biology
Round 2 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	381	400	411
Median	2	380	394	403
Median	Overall	380	396	406
25th %ile	1	380	399	406
25th %ile	2	380	392	400
25th %ile	Overall	380	394	405
75th %ile	1	381	400	411
75th %ile	2	381	394	406
75th %ile	Overall	381	400	411
Minimum	1	380	396	405
Minimum	2	379	392	399
Minimum	Overall	379	392	399
Maximum	1	381	401	424
Maximum	2	384	396	406
Maximum	Overall	384	401	424

Overall		Basic	Proficient	Advanced
	Median	380	396	406
	25th %ile	380	394	405
	75th %ile	381	400	411
	Minimum	379	392	399
	Maximum	384	401	424

Missouri Biology
Round 2 Median Bookmark Summary

Table	Basic	Proficient	Advanced
1	27	71	84
2	25	60.5	77
Overall	25.5	64	81.5

Impact Data

	Below Basic	Basic	Proficient	Advanced
Overall	16.8%	33.7%	24.1%	25.4%

Missouri Biology
Round 3 Bookmark Placements

Table	Panelist	Basic	Proficient	Advanced
1	463	25	71	80
1	464	26	70	82
1	468	25	67	84
1	469	28	72	84
1	471	28	72	80
1	472	29	75	84
2	461	27	68	84
2	462	34	73	84
2	465	24	63	75
2	466	25	75	82
2	467	27	64	80
2	470	22	63	84
Overall		**Basic**	**Proficient**	**Advanced**
	Median	26.5	70.5	83
	25th %ile	25	66.25	80
	75th %ile	28	72.25	84
	Minimum	22	63	75
	Maximum	34	75	84

Missouri Biology
Round 3 Cut Scores

Table	Panelist	Basic	Proficient	Advanced
1	463	380	400	405
1	464	380	399	406
1	468	380	397	411
1	469	381	400	411
1	471	381	400	405
1	472	381	401	411
2	461	381	397	411
2	462	384	401	411
2	465	380	394	401
2	466	380	401	406
2	467	381	396	405
2	470	379	394	411
Overall		**Basic**	**Proficient**	**Advanced**
	Median	381	400	406
	25th %ile	380	397	405
	75th %ile	381	401	411
	Minimum	379	394	401
	Maximum	384	401	411

Missouri Biology
Round 3 Summary of Bookmark Placements

Statistic	Table	Basic	Proficient	Advanced
Median	1	27	71.5	83
Median	2	26	66	83
Median	Overall	26.5	70.5	83
25th %ile	1	25.25	70.25	80.5
25th %ile	2	24.25	63.25	80.5
25th %ile	Overall	25	66.25	80
75th %ile	1	28	72	84
75th %ile	2	27	71.75	84
75th %ile	Overall	28	72.25	84
Minimum	1	25	67	80
Minimum	2	22	63	75
Minimum	Overall	22	63	75
Maximum	1	29	75	84
Maximum	2	34	75	84
Maximum	Overall	34	75	84

Overall		Basic	Proficient	Advanced
	Median	26.5	70.5	83
	25th %ile	25	66.25	80
	75th %ile	28	72.25	84
	Minimum	22	63	75
	Maximum	34	75	84

Missouri Biology
Round 3 Summary of Cut Scores

Statistic	Table	Basic	Proficient	Advanced
Median	1	381	400	406
Median	2	380	396	406
Median	Overall	381	400	406
25th %ile	1	380	400	406
25th %ile	2	380	396	406
25th %ile	Overall	380	397	405
75th %ile	1	381	400	411
75th %ile	2	381	400	411
75th %ile	Overall	381	401	411
Minimum	1	380	397	405
Minimum	2	379	394	401
Minimum	Overall	379	394	401
Maximum	1	381	401	411
Maximum	2	384	401	411
Maximum	Overall	384	401	411

Overall		Basic	Proficient	Advanced
	Median	381	400	406
	25th %ile	380	397	405
	75th %ile	381	401	411
	Minimum	379	394	401
	Maximum	384	401	411

Missouri Biology
Round 3 Median Bookmark Summary

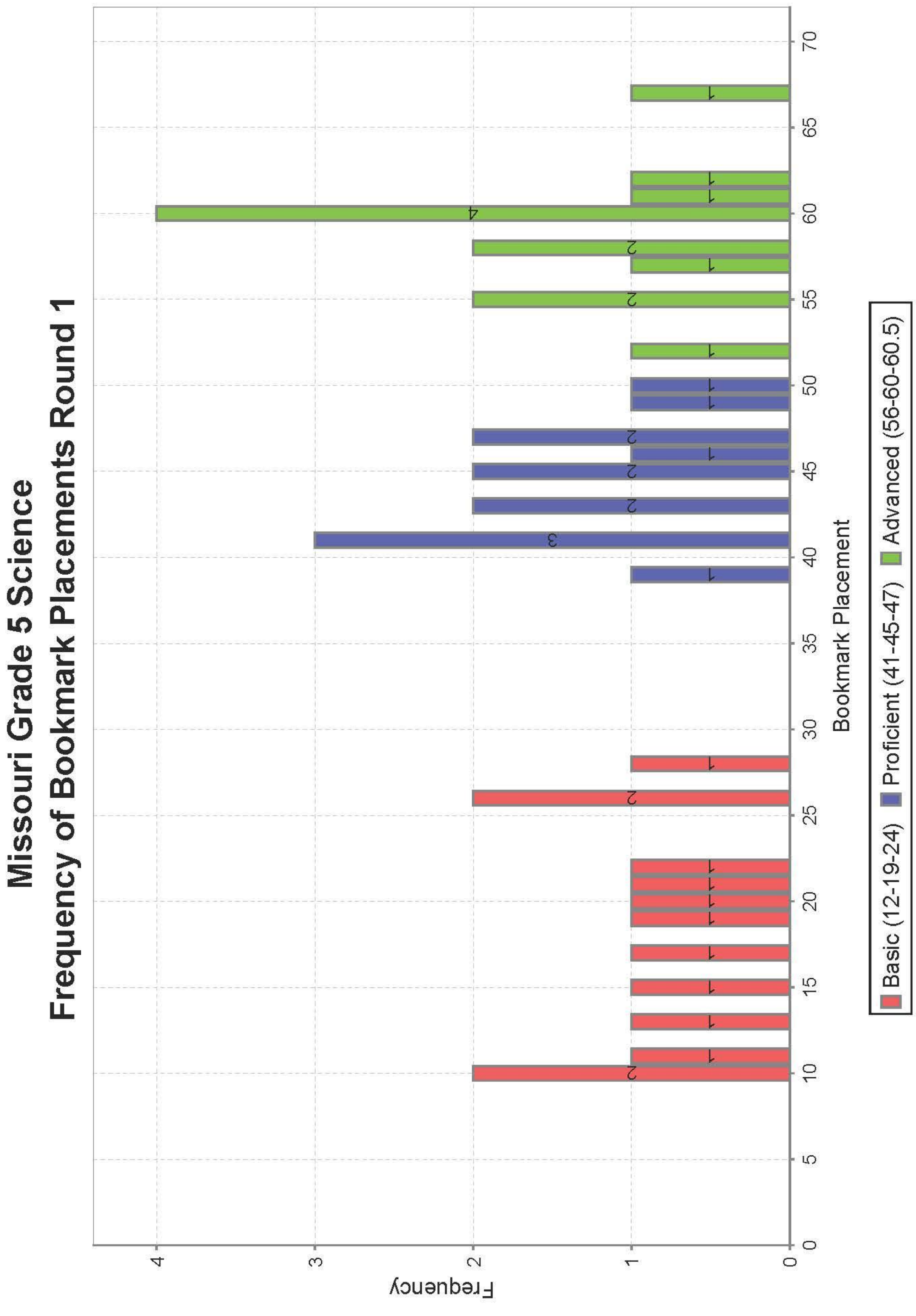
Table	Basic	Proficient	Advanced
1	27	71.5	83
2	26	66	83
Overall	26.5	70.5	83

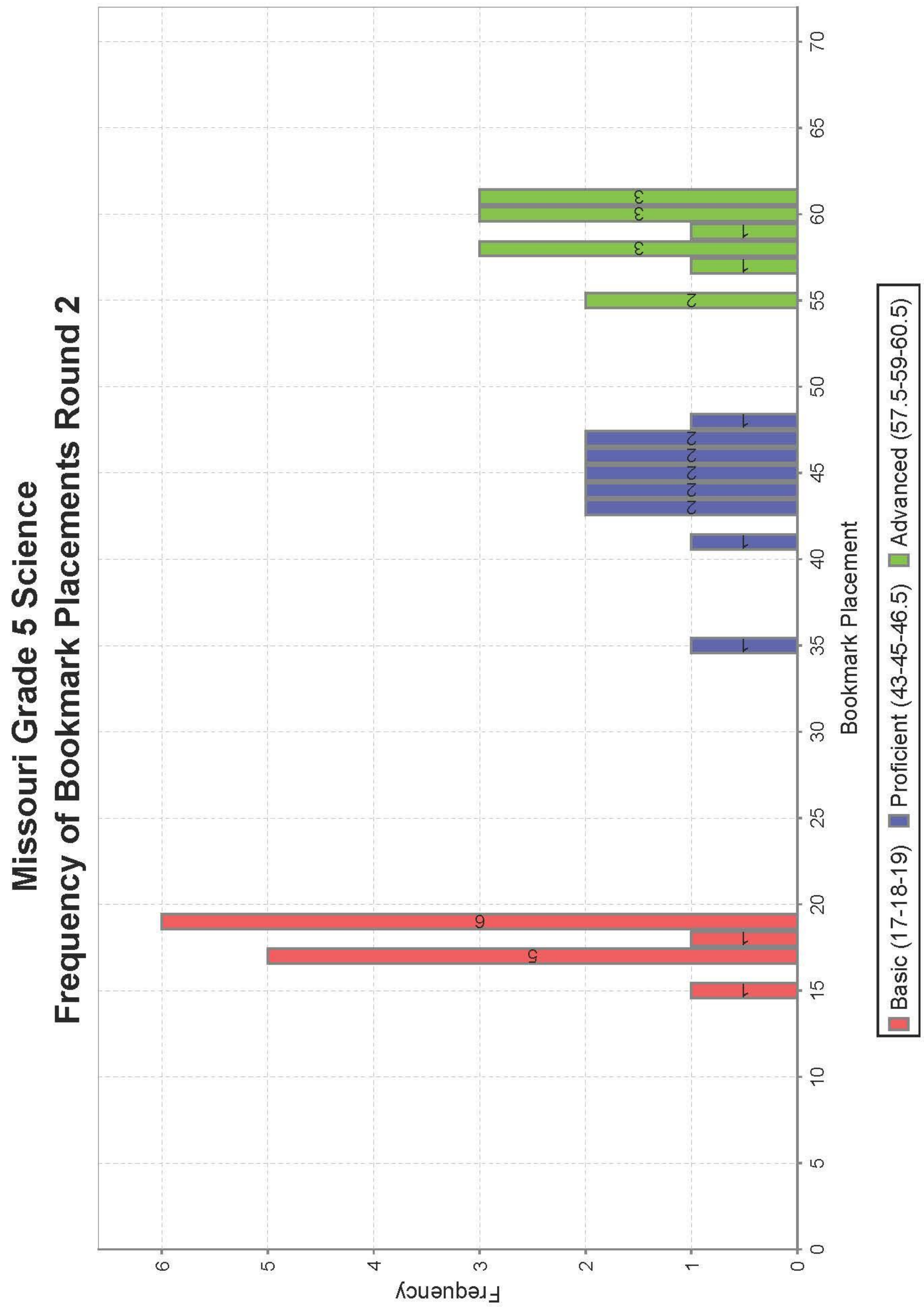
Impact Data

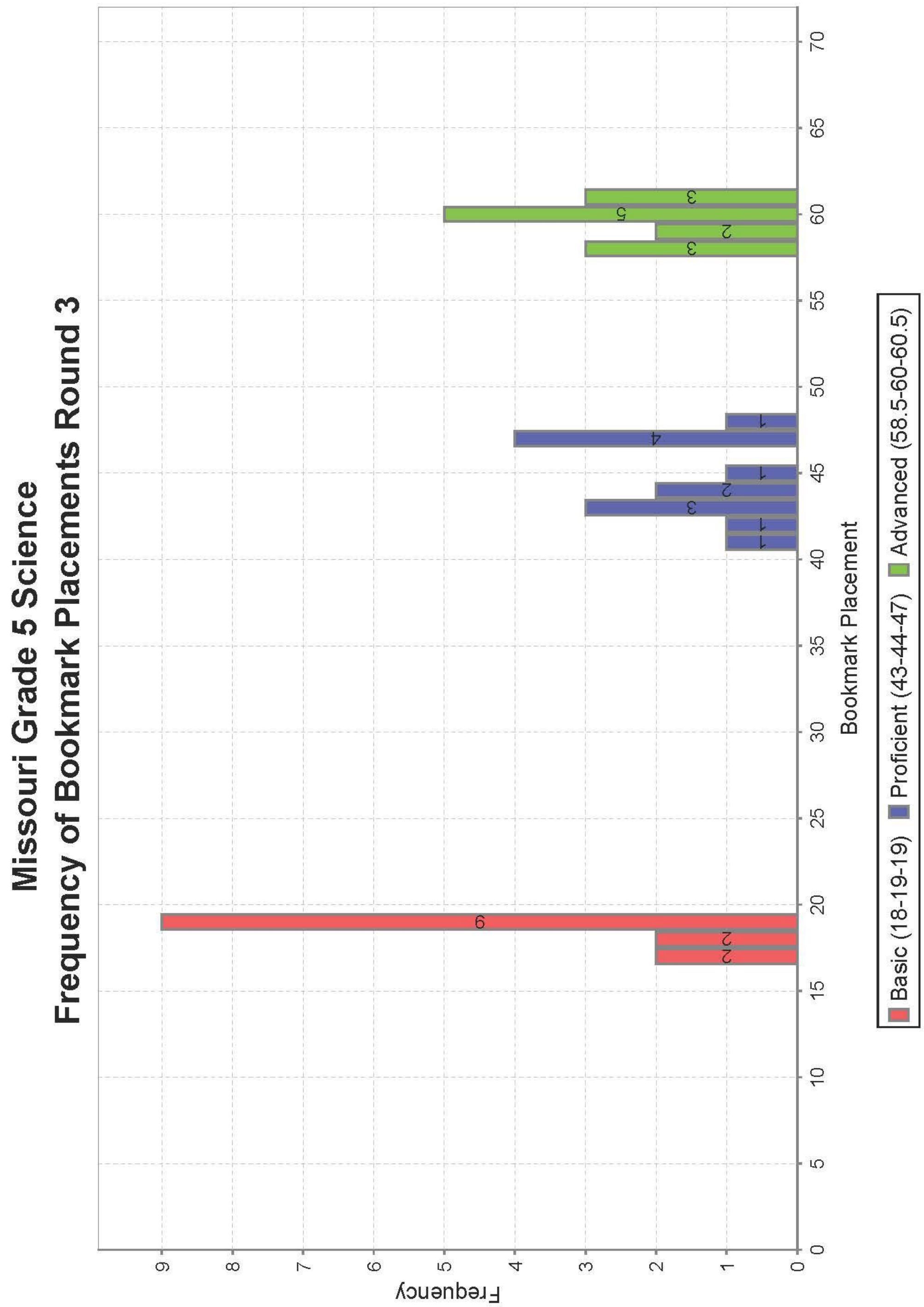
	Below Basic	Basic	Proficient	Advanced
Overall	17.7%	42.9%	14.1%	25.4%

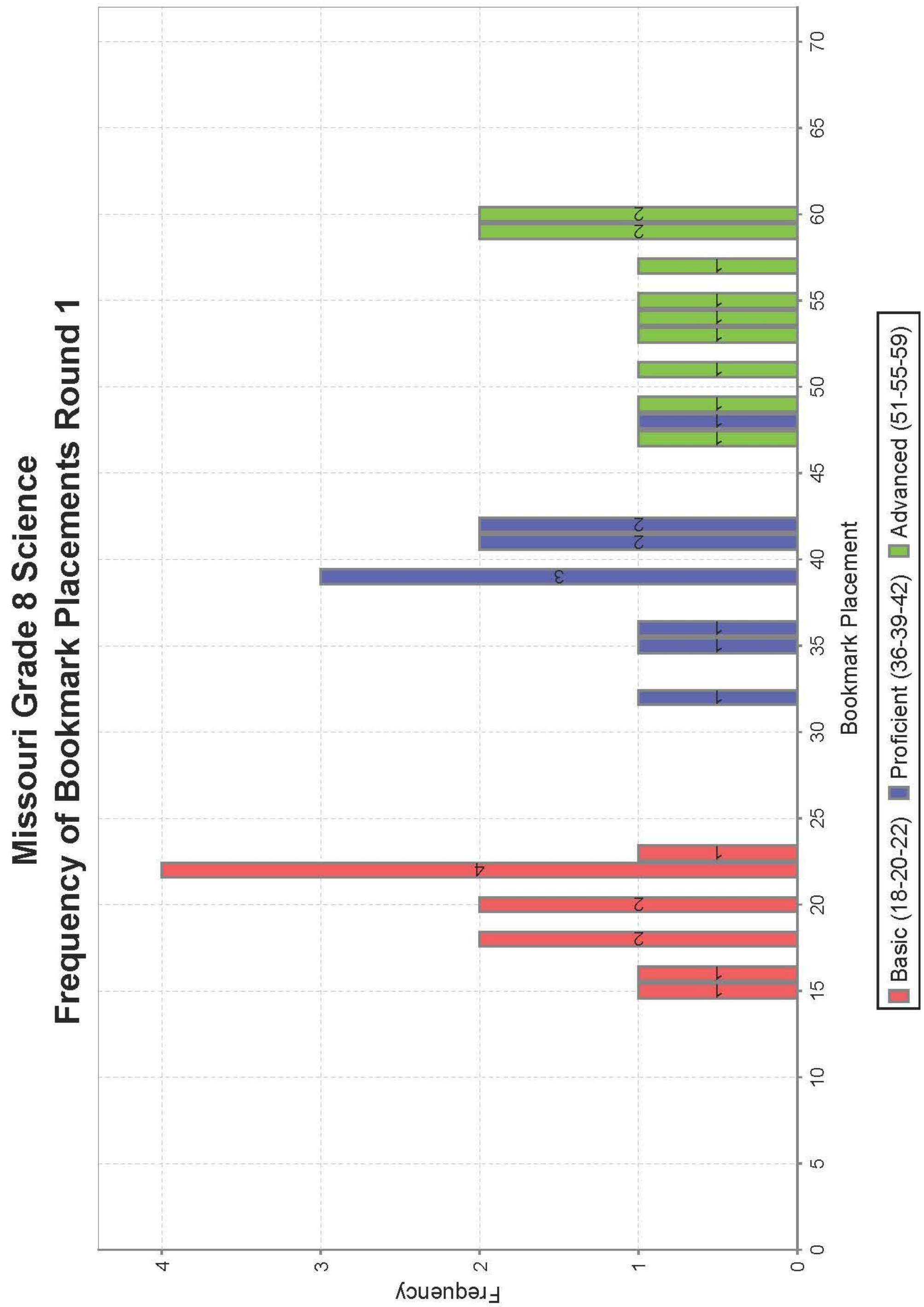
G

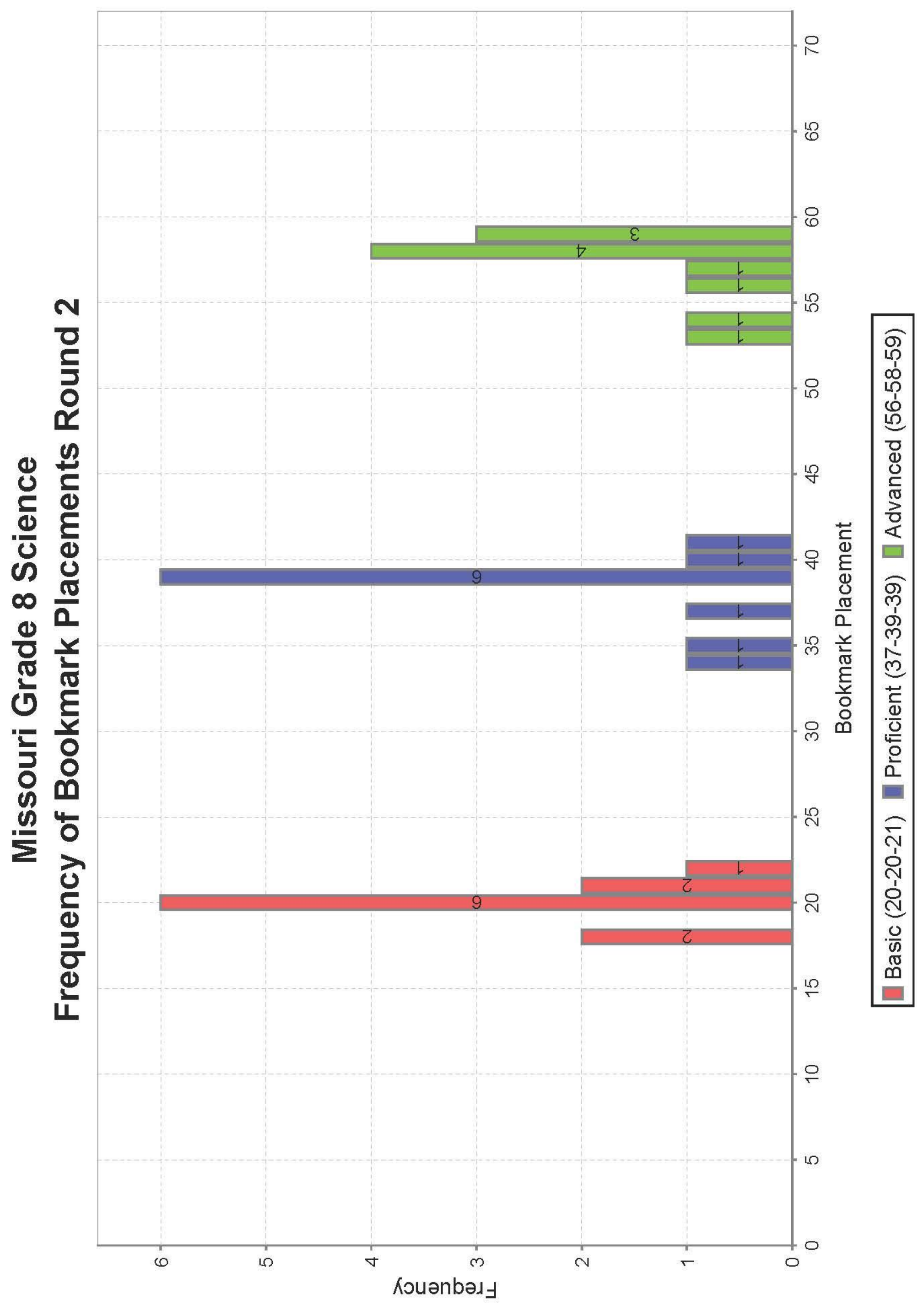
Graphical Representation of Participants' Judgments

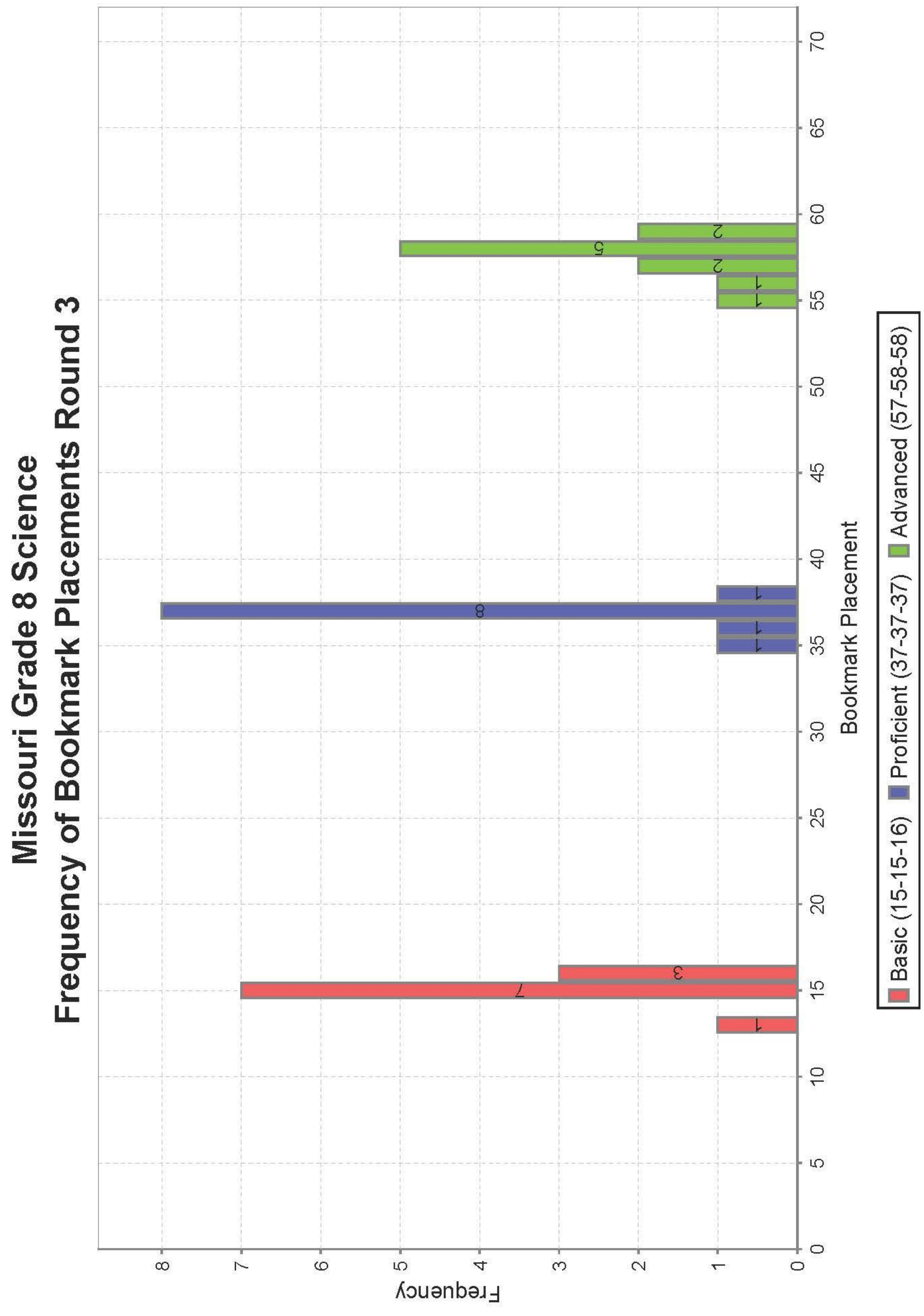

Missouri Grade 5 Science
Frequency of Bookmark Placements Round 1
Frequency
Bookmark Placement
Basic (12-19-24)
Proficient (41-45-47)
Advanced (56-60-60.5)


Missouri Grade 5 Science
Frequency of Bookmark Placements Round 2
Frequency
Bookmark Placement
Basic (17-18-19)
Proficient (43-45-46.5)
Advanced (57.5-59-60.5)


Missouri Grade 5 Science
Frequency of Bookmark Placements Round 3
Frequency
Bookmark Placement
Basic (18-19-19)
Proficient (43-44-47)
Advanced (58.5-60-60.5)

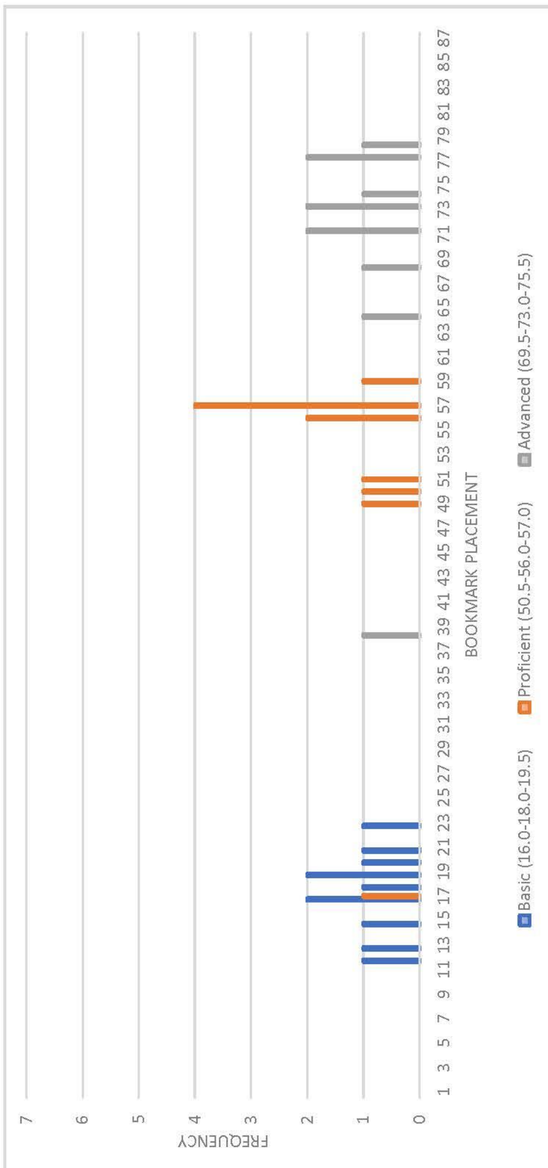

Missouri Grade 8 Science
Frequency of Bookmark Placements Round 1
Frequency
Bookmark Placement
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
0
1
2
3
4
Basic (18-20-22)
Proficient (36-39-42)
Advanced (51-55-59)

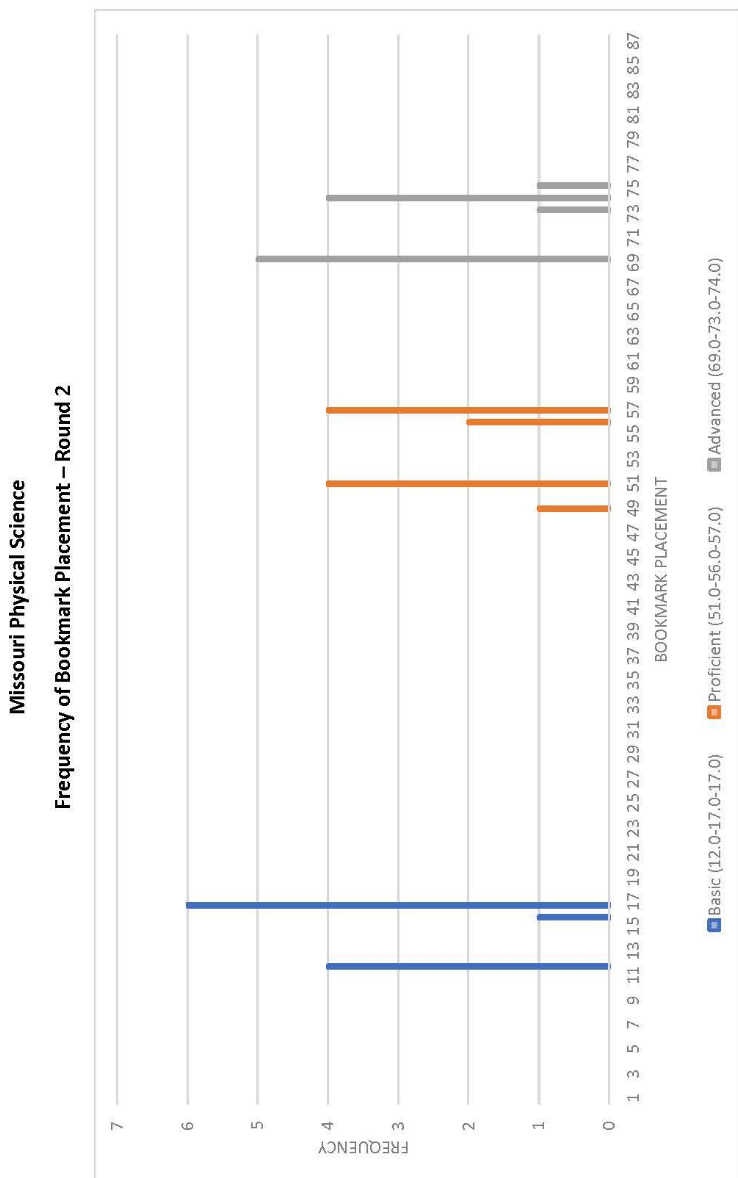

Missouri Grade 8 Science
Frequency of Bookmark Placements Round 2
Frequency
Bookmark Placement
Basic (20-20-21)
Proficient (37-39-39)
Advanced (56-58-59)


Missouri Grade 8 Science
Frequency of Bookmark Placements Round 3
Frequency
Bookmark Placement
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
0
1
2
3
4
5
6
7
8
Basic (15-15-16)
Proficient (37-37-37)
Advanced (57-58-58)

Missouri Physical Science

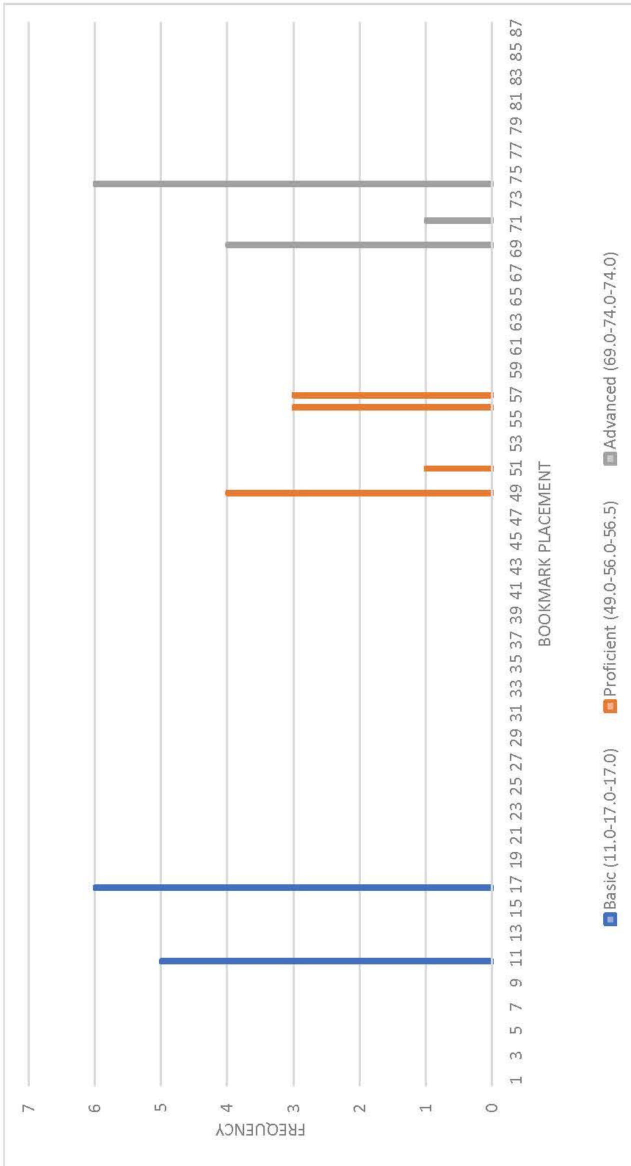
Frequency of Bookmark Placement – Round 1




Missouri Physical Science
Frequency of Bookmark Placement – Round 2
FREQUENCY
BOOKMARK PLACEMENT
Basic (12.0-17.0-17.0)
Proficient (51.0-56.0-57.0)
Advanced (69.0-73.0-74.0)

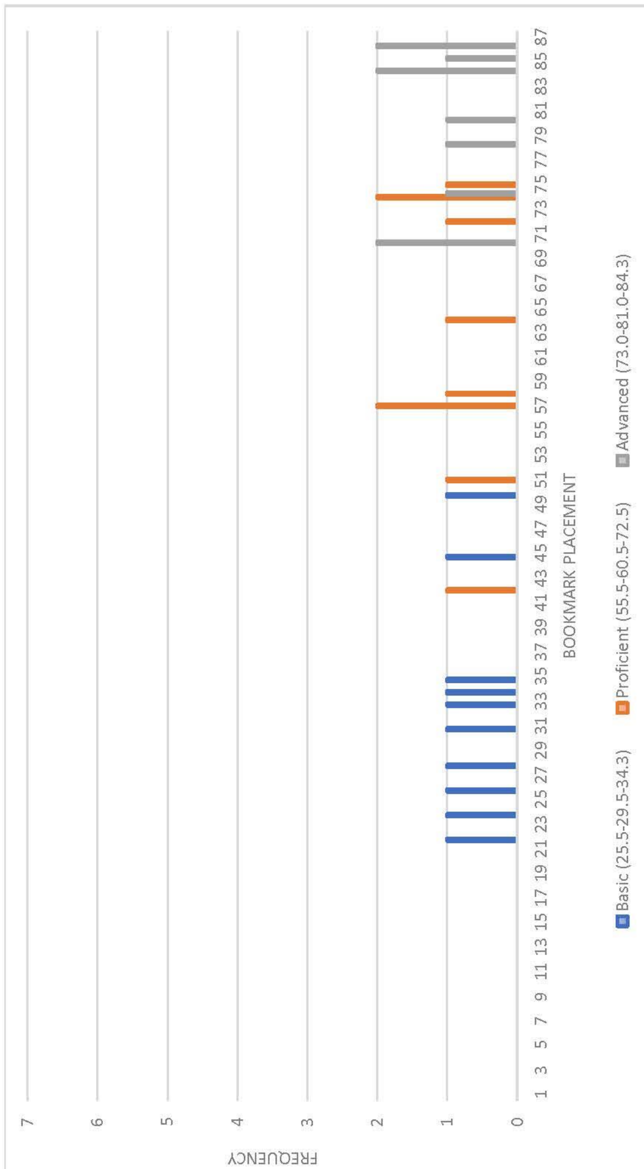
Missouri Physical Science

Frequency of Bookmark Placement – Round 3



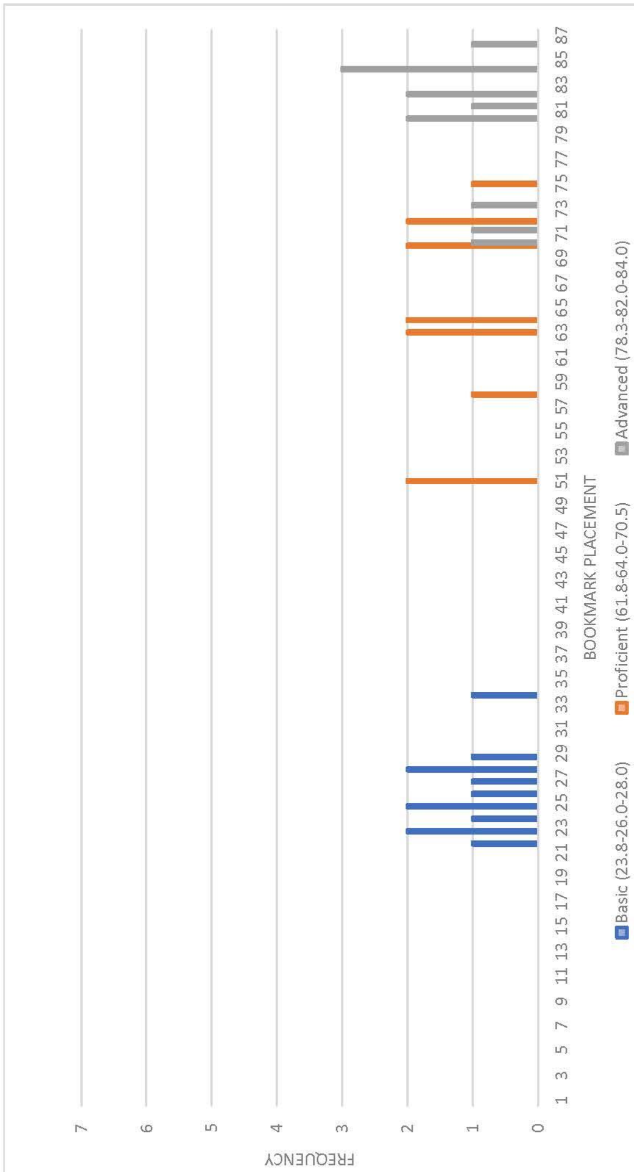
Missouri Biology

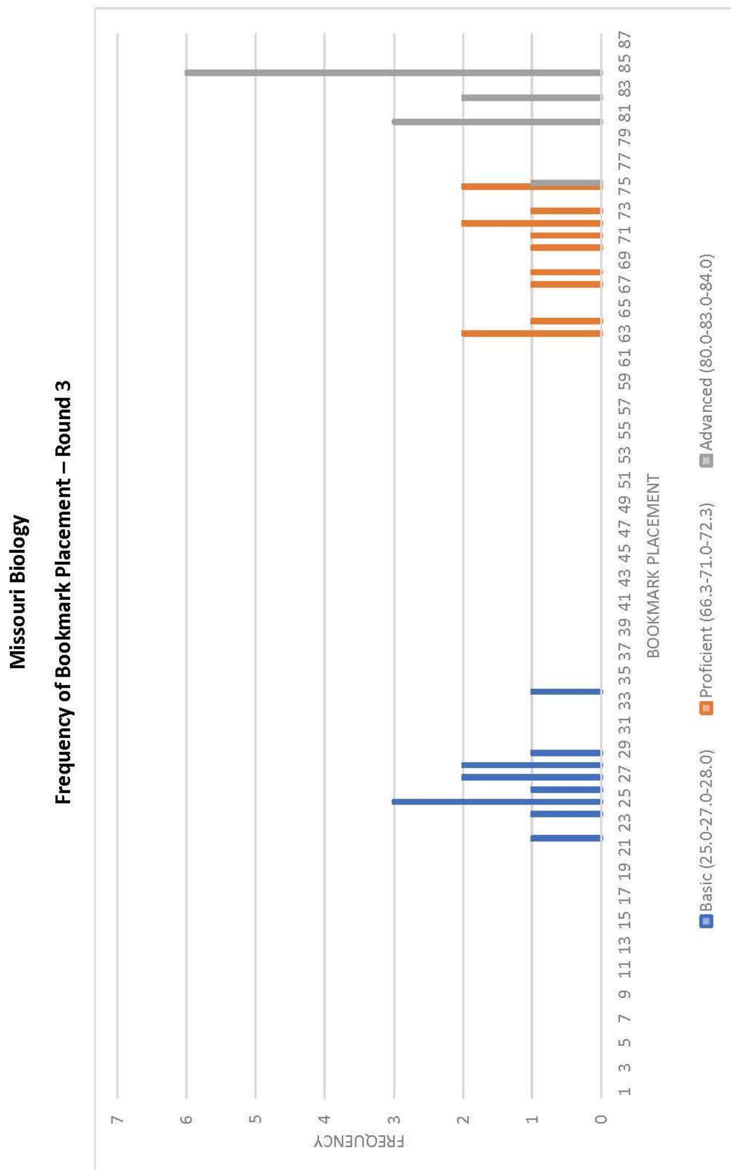
Frequency of Bookmark Placement – Round 1



Missouri Biology

Frequency of Bookmark Placement – Round 2




Missouri Biology
Frequency of Bookmark Placement – Round 3
FREQUENCY
0
1
2
3
4
5
6
7
BOOKMARK PLACEMENT
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63 65 67 69 71 73 75 77 79 81 83 85 87
Basic (25.0-27.0-28.0)
Proficient (66.3-71.0-72.3)
Advanced (80.0-83.0-84.0)

H

Standard Errors Associated with Cut Scores

Missouri Grade 5 Science

Recommended Cut Points* Plus/Minus Selected Standard Errors (SEs) of Measurement

Performance Level	Below Basic	Basic	Proficient	Advanced	
Standard Error (SE) measurement		11.00	11.00	12.00	
Recommended Cut Point* + 3 SE		308	343	380	+ 3 SE
Percent of Students in Each Level	55.2	30.6	12.7	1.5	
Recommended Cut Point* + 2 SE		297	332	368	+ 2 SE
Percent of Students in Each Level	44.5	33.2	18.8	3.5	
Recommended Cut Point* + 1 SE		286	321	356	+ 1 SE
Percent of Students in Each Level	34.5	33.3	24.9	7.3	
Recommended Cut Point*		275	310	344	Recommended Cut Points*
Percent of Students in Each Level	26.0	31.1	29.3	13.5	
Recommended Cut Point* -1 SE		264	299	332	-1 SE
Percent of Students in Each Level	18.8	27.5	31.3	22.3	
Recommended Cut Point* -2 SE		253	288	320	-2 SE
Percent of Students in Each Level	12.9	23.4	30.5	33.2	
Recommended Cut Point* -3 SE		242	277	308	-3 SE
Percent of Students in Each Level	8.1	19.4	27.7	44.8	

* Participants' Large Group Medians

Missouri Grade 8 Science

Recommended Cut Points* Plus/Minus Selected Standard Errors (SEs) of Measurement

Performance Level	Below Basic	Basic	Proficient	Advanced	
Standard Error (SE) measurement		11.00	9.00	9.00	
Recommended Cut Point* + 3 SE		494	528	564	+ 3 SE
Percent of Students in Each Level	40.1	34.2	22.1	3.6	
Recommended Cut Point* + 2 SE		483	519	555	+ 2 SE
Percent of Students in Each Level	30.8	34.7	27.8	6.8	
Recommended Cut Point* + 1 SE		472	510	546	+ 1 SE
Percent of Students in Each Level	22.9	33.2	32.3	11.6	
Recommended Cut Point*		461	501	537	Recommended Cut Points*
Percent of Students in Each Level	16.5	30.4	35.0	18.1	
Recommended Cut Point* -1 SE		450	492	528	-1 SE
Percent of Students in Each Level	11.5	26.7	36.0	25.8	
Recommended Cut Point* -2 SE		439	483	519	-2 SE
Percent of Students in Each Level	7.7	23.0	34.7	34.5	
Recommended Cut Point* -3 SE		428	474	510	-3 SE
Percent of Students in Each Level	4.8	19.4	31.9	43.9	

* Participants' Large Group Medians

Missouri Physical Science EOC

Recommended cut points* Plus/Minus Selected Standard Error of Measurement

Performance Level	Below Basic	Basic	Proficient	Advanced	Performance Level
Standard Error (SE) of Measurement		5	5	6	
Recommended Cut Point + 3SE		397	415	435	+ 3SE
% of students in each level	51.48%	36.09%	10.90%	1.53%	
Recommended Cut Point + 2SE		392	410	429	+ 2SE
% of students in each level	32.58%	48.99%	14.27%	4.17%	
Recommended Cut Point + 1SE		387	405	423	+ 1SE
% of students in each level	24.96%	46.14%	19.97%	8.93%	
Recommended Cut Point		**382**	**400**	**417**	**Recommended Cut point**
% of students in each level	**16.77%**	**43.77%**	**24.49%**	**14.97%**	
Recommended Cut Point - 1SE		377	395	410	- 1 SE
% of students in each level	10.79%	39.73%	25.99%	23.50%	
Recommended Cut point - 2SE		371	390	404	- 2SE
% of students in each level	5.75%	32.53%	25.72%	35.99%	
Recommended Cut point - 3SE		366	385	398	- 3SE
% of students in each level	0.85%	21.20%	33.00%	44.95%	

* Participants large group medians

Missouri Biology EOC

Recommended cut points* Plus/Minus Selected Standard Error of Measurement

Performance Level	Below Basic	Basic	Proficient	Advanced	Performance Level
Standard Error (SE) of Measurement		4	4	5	
Recommended Cut Point + 3SE		393	413	425	+ 3SE
Percent of students in each level	43.03%	45.22%	9.33%	2.42%	
Recommended Cut Point + 2SE		389	409	421	+ 2SE
Percent of students in each level	32.58%	48.99%	14.27%	4.17%	
Recommended Cut Point + 1SE		385	404	416	+ 1SE
Percent of students in each level	24.96%	46.14%	19.97%	8.93%	
Recommended Cut Point		**381**	**400**	**411**	**Recommended Cut point**
Percent of students in each level	**16.77%**	**43.77%**	**24.49%**	**14.97%**	
Recommended Cut Point - 1SE		377	396	406	- 1 SE
Percent of students in each level	10.79%	39.73%	25.99%	23.50%	
Recommended Cut point - 2SE		373	391	402	- 2SE
Percent of students in each level	5.75%	32.53%	25.72%	35.99%	
Recommended Cut point - 3SE		369	387	397	- 3SE
Percent of students in each level	2.59%	26.32%	24.90%	46.19%	

* Participants large group medians

I
Participant Evaluations of the Workshop

Missouri MAP Science 2019 Standard Setting Evaluation

The purpose of this evaluation is to help document the process used to recommend performance standards for Missouri's science tests. Your opinions and comments are important, as they provide a basis for judging the quality of this process.

Please do not put your name on this form. While we need the information to examine the success of the various steps in the process, we want your comments to remain anonymous. This information will be reported only in the aggregate. **When you have completed the evaluation, please give it to a facilitator. *Thank you!***

	Part 1: ABOUT THE STANDARD SETTING Please *consider the statements below* and mark the level of agreement or disagreement you have with each statement. Please bubble ***only one*** of the five options for each statement.	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Training & Process	1. I understood the purpose of this workshop.	O	O	O	O	O
	2. The pre-workshop assignment was helpful for me to engage in the workshop activities.	O	O	O	O	O
	3. I understood the content measured by the assessment I reviewed at the standard setting.	O	O	O	O	O
	4. I understood how the assessment was administered.	O	O	O	O	O
	5. I understood the difference between range PLDs and threshold student descriptors.	O	O	O	O	O
	6. The PLDs were clear enough for me to describe the threshold students.	O	O	O	O	O
	7. Before Round 1 began, I was comfortable with the Bookmark Procedure.	O	O	O	O	O
	8. I understood how to use the item map.	O	O	O	O	O
	9. I understood the ordered item booklet (and passage booklet).	O	O	O	O	O
	10. I understood how to place my bookmarks.	O	O	O	O	O
	11. I understood the room-level data that was presented between the rounds.	O	O	O	O	O
	12. I understood what the benchmarks were and how I should consider them.	O	O	O	O	O
	13. I understood the impact data that were presented.	O	O	O	O	O
	14. By the end of the workshop, I was comfortable with the Bookmark Procedure.	O	O	O	O	O
	15. The instructions provided in the training materials were clear.	O	O	O	O	O
	16. The instructions provided during the opening training session were clear.	O	O	O	O	O
	17. The instructions provided by the facilitators in my breakout room were clear.	O	O	O	O	O
	18. Overall, I believe my opinions were considered and valued by my group.	O	O	O	O	O
	19. My group's work was reflected in the presentation of recommendations across grades.	O	O	O	O	O
	20. Overall, I valued the workshop as a professional development experience.	O	O	O	O	O
	21. This process will lead to defensible performance standards for the test.	O	O	O	O	O

	Please indicate your opinion regarding the **usefulness** of the following materials used. Please bubble ***only one*** of the four options for each material.	Not Useful	Somewhat Useful	Useful	Very Useful
Usefulness	22. Taking and discussing the test before placing a bookmark.	O	O	O	O
	23. Describing the three categories of threshold students.	O	O	O	O
	24. Reviewing the ordered item booklet and passage booklet before placing a bookmark.	O	O	O	O
	25. The item map.	O	O	O	O
	26. The practice activities on making bookmark placements.	O	O	O	O
	27. Table-level discussion.	O	O	O	O
	28. Learning about the other tables' discussions.	O	O	O	O
	29. Large-group feedback and discussion.	O	O	O	O
	30. The benchmarks.	O	O	O	O
	31. The percent of students in each performance level (the impact data).	O	O	O	O
	32. The across-grade presentation of impact data.	O	O	O	O

	Please indicate your opinion regarding how **influential** the following were when you placed your bookmarks. Please bubble ***only one*** of the four options for each material.	Not Influential	Somewhat Influential	Influential	Very Influential
Influence	33. The performance level descriptors (PLDs).	O	O	O	O
	34. The descriptions of the threshold students.	O	O	O	O
	35. My perception of the difficulty of the items.	O	O	O	O
	36. My experiences with students.	O	O	O	O
	37. Discussion at my table.	O	O	O	O
	38. Discussion within my group.	O	O	O	O
	39. The bookmark placements of other participants.	O	O	O	O
	40. The percent of students in each performance level (the impact data).	O	O	O	O
	41. The benchmarks.	O	O	O	O
	42. My sense of what a student needs to know in order to be *Basic*.	O	O	O	O
	43. My sense of what a student needs to know in order to be *Proficient*.	O	O	O	O
	44. My sense of what a student needs to know in order to be *Advanced*.	O	O	O	O

	Please indicate your opinion regarding **how much time** you were given to complete the following activities. Please bubble ***only one*** of the three options for each activity.	Too Little Time	About Right	Too Much Time
Time	45. The general session.	O	O	O
	46. Reviewing the test.	O	O	O
	47. Reviewing the PLDs.	O	O	O
	48. Describing the threshold students.	O	O	O
	49. Reviewing the ordered item booklet (OIB).	O	O	O
	50. Reviewing the item map.	O	O	O
	51. Training in the Bookmark Procedure.	O	O	O
	52. Practice activities making bookmark placements.	O	O	O
	53. Table discussions after Round 1.	O	O	O
	54. Table discussions after Round 2.	O	O	O
	55. Group discussion after Round 2.	O	O	O

Please indicate your opinion regarding whether you feel the group's final, recommended cut scores were **too low, about right, or too high** for each cut score. Please bubble ***only one*** of the three options for each cut score.	Too Low	About Right	Too High
56. *Basic* cut score	O	O	O
57. *Proficient* cut score	O	O	O
58. *Advanced* cut score	O	O	O

Please consider the process that the group followed to achieve the final, recommended cut scores. Overall, **how comfortable are you with the group's recommendations**? Please indicate your opinion for each cut score. Please bubble ***only one*** of the five options for each cut score.	Very Uncomfortable	Somewhat Uncomfortable	Neutral	Somewhat Comfortable	Very Comfortable
59. *Basic* cut score	O	O	O	O	O
60. *Proficient* cut score	O	O	O	O	O
61. *Advanced* cut score	O	O	O	O	O

Part 2: ABOUT YOU

62. In which group did you work?
- O Grade 5 Science
- O Grade 8 Science
- O Physical Science
- O Biology

63. What is your current assignment?
- O Classroom teacher
- O Educator, non-teacher
- O Higher education
- O Other, (please describe): ______________

64. How many years, in total, have you been teaching?
- O Fewer than 5 years
- O 5–10 years
- O 11–15 years
- O 16–20 years
- O 21–25 years
- O More than 25 years

65. What is your gender?
- O Female
- O Male

66. What is your ethnicity?
- O American Indian/Alaska Native
- O Asian
- O Hawaiian or Pacific Islander
- O Black
- O Hispanic
- O Mixed (Two or more races)
- O Caucasian

67. What is your highest level of education?
- O High school diploma
- O Associate's degree
- O Bachelor's degree
- O Bachelor's degree + Hours
- O Master's degree
- O Master's degree + Hours
- O Doctoral degree

68. Which of these groups do you have experience teaching?
- O Special education (in a self-contained classroom)
- O Special education (in a mainstream classroom)
- O English language learners
- O Gifted and talented
- O Vocational education

69. What percent of students qualify for free/reduced-price meals at your school?
- O 0 – 25%
- O 26 – 50%
- O 51 – 75%
- O 76% or higher

Part 3: YOUR TURN

In this box, please feel free to add comments about any of your responses, make suggestions to improve future workshops, or tell us what you liked and did not like about this workshop. ***Thank you!***

Grade 5 Science

1. I understood the purpose of this workshop.

Response	Frequency	Percent	Mean: 4.69
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	30.77	
Strongly Agree	9	69.23	

2. The pre-workshop assignment was helpful for me to engage in the workshop activities.

Response	Frequency	Percent	Mean: 4.00
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	3	23.08	
Agree	6	46.15	
Strongly Agree	3	23.08	
No Response	1	7.69	

3. I understood the content measured by the assessment I reviewed at the standard setting.

Response	Frequency	Percent	Mean: 4.38
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	8	61.54	
Strongly Agree	5	38.46	

4. I understood how the assessment was administered.

Response	Frequency	Percent	Mean: 4.46
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	53.85	
Strongly Agree	6	46.15	

5. I understood the difference between range PLDs and threshold student descriptors.

Response	Frequency	Percent	Mean: 4.54
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	46.15	
Strongly Agree	7	53.85	

6. The PLDs were clear enough for me to describe the threshold students.

Response	Frequency	Percent	Mean: 4.38
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	8	61.54	
Strongly Agree	5	38.46	

7. Before Round 1 began, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 4.23
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	7.69	
Agree	8	61.54	
Strongly Agree	4	30.77	

8. I understood how to use the item map.

Response	Frequency	Percent	Mean: 4.31
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	7.69	
Agree	7	53.85	
Strongly Agree	5	38.46	

9. I understood the ordered item booklet (and passage booklet).

Response	Frequency	Percent	Mean: 4.62
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	38.46	
Strongly Agree	8	61.54	

10. I understood how to place my bookmarks.

Response	Frequency	Percent	Mean: 4.38
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	8	61.54	
Strongly Agree	5	38.46	

11. I understood the room-level data that was presented between the rounds.

Response	Frequency	Percent	Mean: 4.69
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	30.77	
Strongly Agree	9	69.23	

12. I understood what the benchmarks were and how I should consider them.

Response	Frequency	Percent	Mean: 4.69
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	30.77	
Strongly Agree	9	69.23	

13. I understood the impact data that were presented.

Response	Frequency	Percent	Mean: 4.69
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	30.77	
Strongly Agree	9	69.23	

14. By the end of the workshop, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 4.77
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	23.08	
Strongly Agree	10	76.92	

15. The instructions provided in the training materials were clear.

Response	Frequency	Percent	Mean: 4.46
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	53.85	
Strongly Agree	6	46.15	

16. The instructions provided during the opening training session were clear.

Response	Frequency	Percent	Mean: 4.38
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	8	61.54	
Strongly Agree	5	38.46	

17. The instructions provided by the facilitators in my breakout room were clear.

Response	Frequency	Percent	Mean: 4.62
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	38.46	
Strongly Agree	8	61.54	

18. Overall, I believe my opinions were considered and valued by my group.

Response	Frequency	Percent	Mean: 4.54
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	46.15	
Strongly Agree	7	53.85	

19. My group's work was reflected in the presentation of recommendations across grades.

Response	Frequency	Percent	Mean: 4.67
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	30.77	
Strongly Agree	8	61.54	
No Response	1	7.69	

20. Overall, I valued the workshop as a professional development experience.

Response	Frequency	Percent	Mean: 4.62
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	38.46	
Strongly Agree	8	61.54	

21. This process will lead to defensible performance standards for the test.

Response	Frequency	Percent	Mean: 4.31
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	7.69	
Agree	7	53.85	
Strongly Agree	5	38.46	

22. Taking and discussing the test before placing a bookmark.

Response	Frequency	Percent	Mean: 3.46
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	7	53.85	
Very Useful	6	46.15	

23. Describing the three categories of threshold students.

Response	Frequency	Percent	Mean: 3.69
Not Useful	0	0.00	
Somewhat Useful	1	7.69	
Useful	2	15.38	
Very Useful	10	76.92	

24. Reviewing the ordered item booklet and passage booklet before placing a bookmark.

Response	Frequency	Percent	Mean: 3.54
Not Useful	0	0.00	
Somewhat Useful	2	15.38	
Useful	2	15.38	
Very Useful	9	69.23	

25. The item map.

Response	Frequency	Percent	Mean: 3.31
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	9	69.23	
Very Useful	4	30.77	

26. The practice activities on making bookmark placements.

Response	Frequency	Percent	Mean: 3.31
Not Useful	0	0.00	
Somewhat Useful	1	7.69	
Useful	7	53.85	
Very Useful	5	38.46	

27. Table-level discussion.

Response	Frequency	Percent	Mean: 3.62
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	5	38.46	
Very Useful	8	61.54	

28. Learning about the other tables' discussions.

Response	Frequency	Percent	Mean: 3.46
Not Useful	1	7.69	
Somewhat Useful	0	0.00	
Useful	4	30.77	
Very Useful	8	61.54	

29. Large-group feedback and discussion.

Response	Frequency	Percent	Mean: 3.46
Not Useful	0	0.00	
Somewhat Useful	2	15.38	
Useful	3	23.08	
Very Useful	8	61.54	

30. The benchmarks.

Response	Frequency	Percent	Mean: 3.31
Not Useful	0	0.00	
Somewhat Useful	1	7.69	
Useful	7	53.85	
Very Useful	5	38.46	

31. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 3.31
Not Useful	1	7.69	
Somewhat Useful	1	7.69	
Useful	4	30.77	
Very Useful	7	53.85	

32. The across-grade presentation of impact data.

Response	Frequency	Percent	Mean: 3.25
Not Useful	0	0.00	
Somewhat Useful	2	15.38	
Useful	5	38.46	
Very Useful	5	38.46	
No Response	1	7.69	

33. The performance level descriptors (PLDs).

Response	Frequency	Percent	Mean: 3.38
Not Influential	0	0.00	
Somewhat Influential	1	7.69	
Influential	6	46.15	
Very Influential	6	46.15	

34. The descriptions of the threshold students.

Response	Frequency	Percent	Mean: 3.54
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	6	46.15	
Very Influential	7	53.85	

35. My perception of the difficulty of the items.

Response	Frequency	Percent	Mean: 2.92
Not Influential	0	0.00	
Somewhat Influential	2	15.38	
Influential	10	76.92	
Very Influential	1	7.69	

36. My experiences with students.

Response	Frequency	Percent	Mean: 2.69
Not Influential	1	7.69	
Somewhat Influential	6	46.15	
Influential	2	15.38	
Very Influential	4	30.77	

37. Discussion at my table.

Response	Frequency	Percent	Mean: 3.38
Not Influential	0	0.00	
Somewhat Influential	1	7.69	
Influential	6	46.15	
Very Influential	6	46.15	

38. Discussion within my group.

Response	Frequency	Percent	Mean: 3.31
Not Influential	0	0.00	
Somewhat Influential	1	7.69	
Influential	7	53.85	
Very Influential	5	38.46	

39. The bookmark placements of other participants.

Response	Frequency	Percent	Mean: 2.46
Not Influential	0	0.00	
Somewhat Influential	7	53.85	
Influential	6	46.15	
Very Influential	0	0.00	

40. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 2.46
Not Influential	3	23.08	
Somewhat Influential	3	23.08	
Influential	5	38.46	
Very Influential	2	15.38	

41. The benchmarks.

Response	Frequency	Percent	Mean: 2.92
Not Influential	0	0.00	
Somewhat Influential	3	23.08	
Influential	8	61.54	
Very Influential	2	15.38	

42. My sense of what a student needs to know in order to be Basic.

Response	Frequency	Percent	Mean: 3.00
Not Influential	1	7.69	
Somewhat Influential	2	15.38	
Influential	6	46.15	
Very Influential	4	30.77	

43. My sense of what a student needs to know in order to be Proficient.

Response	Frequency	Percent	Mean: 3.00
Not Influential	1	7.69	
Somewhat Influential	2	15.38	
Influential	6	46.15	
Very Influential	4	30.77	

44. My sense of what a student needs to know in order to be Advanced.

Response	Frequency	Percent	Mean: 3.08
Not Influential	1	7.69	
Somewhat Influential	1	7.69	
Influential	7	53.85	
Very Influential	4	30.77	

45. The general session.

Response	Frequency	Percent	Mean: 2.31
Too Little Time	0	0.00	
About Right	9	69.23	
Too Much Time	4	30.77	

46. Reviewing the test.

Response	Frequency	Percent	Mean: 1.92
Too Little Time	1	7.69	
About Right	12	92.31	
Too Much Time	0	0.00	

47. Reviewing the PLDs.

Response	Frequency	Percent	Mean: 2.08
Too Little Time	0	0.00	
About Right	12	92.31	
Too Much Time	1	7.69	

48. Describing the threshold students.

Response	Frequency	Percent	Mean: 2.08
Too Little Time	0	0.00	
About Right	12	92.31	
Too Much Time	1	7.69	

49. Reviewing the ordered item booklet (OIB).

Response	Frequency	Percent	Mean: 1.77
Too Little Time	3	23.08	
About Right	10	76.92	
Too Much Time	0	0.00	

50. Reviewing the item map.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	13	100.00	
Too Much Time	0	0.00	

51. Training in the Bookmark Procedure.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	1	7.69	
About Right	11	84.62	
Too Much Time	1	7.69	

52. Practice activities making bookmark placements.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	13	100.00	
Too Much Time	0	0.00	

53. Table discussions after Round 1.

Response	Frequency	Percent	Mean: 2.08
Too Little Time	0	0.00	
About Right	12	92.31	
Too Much Time	1	7.69	

54. Table discussions after Round 2.

Response	Frequency	Percent	Mean: 2.08
Too Little Time	0	0.00	
About Right	12	92.31	
Too Much Time	1	7.69	

55. Group discussion after Round 2.

Response	Frequency	Percent	Mean: 2.08
Too Little Time	0	0.00	
About Right	12	92.31	
Too Much Time	1	7.69	

56. Basic cut score

Response	Frequency	Percent	Mean: 2.15
Too Low	0	0.00	
About Right	11	84.62	
Too High	2	15.38	

57. Proficient cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	13	100.00	
Too High	0	0.00	

58. Advanced cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	13	100.00	
Too High	0	0.00	

59. Basic cut score

Response	Frequency	Percent	Mean: 4.31
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	2	15.38	
Neutral	0	0.00	
Somewhat Comfortable	3	23.08	
Very Comfortable	8	61.54	

60. Proficient cut score

Response	Frequency	Percent	Mean: 4.38
Very Uncomfortable	1	7.69	
Somewhat Uncomfortable	0	0.00	
Neutral	0	0.00	
Somewhat Comfortable	4	30.77	
Very Comfortable	8	61.54	

61. Advanced cut score

Response	Frequency	Percent	Mean: 4.38
Very Uncomfortable	1	7.69	
Somewhat Uncomfortable	0	0.00	
Neutral	0	0.00	
Somewhat Comfortable	4	30.77	
Very Comfortable	8	61.54	

62. In which group did you work?

Response	Frequency	Percent	Mean: 1.00
Grade 5 Science	13	100.00	
Grade 8 Science	0	0.00	
Physical Science	0	0.00	
Biology	0	0.00	

63. What is your current assignment?

Response	Frequency	Percent	Mean: 2.00
Classroom teacher	6	46.15	
Educator non-teacher	4	30.77	
Higher education	0	0.00	
Other	3	23.08	

64. How many years, in total, have you been teaching?

Response	Frequency	Percent	Mean: 3.54
Fewer than 5 years	0	0.00	
5-10 years	4	30.77	
11-15 years	4	30.77	
16-20 years	1	7.69	
21-25 years	2	15.38	
More than 25 years	2	15.38	

65. What is your gender?

Response	Frequency	Percent	Mean: 1.15
Female	11	84.62	
Male	2	15.38	

66. What is your ethnicity?

Response	Frequency	Percent	Mean: 6.77
American Indian/Alaska Native	0	0.00	
Asian	0	0.00	
Hawaiian or Pacific Islander	0	0.00	
Black	1	7.69	
Hispanic	0	0.00	
Mixed -Two or more races	0	0.00	
Caucasian	12	92.31	

67. What is your highest level of education?

Response	Frequency	Percent	Mean: 5.38
High school diploma	0	0.00	
Associate's degree	0	0.00	
Bachelor's degree	0	0.00	
Bachelor's degree + Hours	4	30.77	
Master's degree	0	0.00	
Master's degree + Hours	9	69.23	
Doctoral degree	0	0.00	

68. Which of these groups do you have experience teaching?

Response	Frequency	Percent	Mean: -
Special education -in a self-contained classroom	1	7.69	
Special education -in a mainstream classroom	7	53.85	
English language learners	4	30.77	
Gifted and talented	1	7.69	
Vocational education	0	0.00	
No Response	4	30.77	

69. What percent of students qualify for free/reduced-price meals at your school?

Response	Frequency	Percent	Mean: 3.00
0-25%	0	0.00	
26-50%	3	23.08	
51-75%	7	53.85	
76% or higher	3	23.08	

Grade 8 Science

1. I understood the purpose of this workshop.

Response	Frequency	Percent	Mean: 4.73
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	27.27	
Strongly Agree	8	72.73	

2. The pre-workshop assignment was helpful for me to engage in the workshop activities.

Response	Frequency	Percent	Mean: 4.00
Strongly Disagree	0	0.00	
Disagree	2	18.18	
Neutral	1	9.09	
Agree	3	27.27	
Strongly Agree	5	45.45	

3. I understood the content measured by the assessment I reviewed at the standard setting.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

4. I understood how the assessment was administered.

Response	Frequency	Percent	Mean: 4.45
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	5	45.45	

5. I understood the difference between range PLDs and threshold student descriptors.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

6. The PLDs were clear enough for me to describe the threshold students.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

7. Before Round 1 began, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 3.82
Strongly Disagree	1	9.09	
Disagree	1	9.09	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	3	27.27	

8. I understood how to use the item map.

Response	Frequency	Percent	Mean: 4.27
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	6	54.55	
Strongly Agree	4	36.36	

9. I understood the ordered item booklet (and passage booklet).

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

10. I understood how to place my bookmarks.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

11. I understood the room-level data that was presented between the rounds.

Response	Frequency	Percent	Mean: 4.45
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	5	45.45	

12. I understood what the benchmarks were and how I should consider them.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	5	45.45	
Strongly Agree	5	45.45	

13. I understood the impact data that were presented.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	6	54.55	

14. By the end of the workshop, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 4.73
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	27.27	
Strongly Agree	8	72.73	

15. The instructions provided in the training materials were clear.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	6	54.55	

16. The instructions provided during the opening training session were clear.

Response	Frequency	Percent	Mean: 4.45
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	5	45.45	

17. The instructions provided by the facilitators in my breakout room were clear.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

18. Overall, I believe my opinions were considered and valued by my group.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

19. My group's work was reflected in the presentation of recommendations across grades.

Response	Frequency	Percent	Mean: 4.27
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	6	54.55	
Strongly Agree	4	36.36	

20. Overall, I valued the workshop as a professional development experience.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	6	54.55	

21. This process will lead to defensible performance standards for the test.

Response	Frequency	Percent	Mean: 4.36
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	63.64	
Strongly Agree	4	36.36	

22. Taking and discussing the test before placing a bookmark.

Response	Frequency	Percent	Mean: 3.55
Not Useful	0	0.00	
Somewhat Useful	1	9.09	
Useful	3	27.27	
Very Useful	7	63.64	

23. Describing the three categories of threshold students.

Response	Frequency	Percent	Mean: 3.64
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	4	36.36	
Very Useful	7	63.64	

24. Reviewing the ordered item booklet and passage booklet before placing a bookmark.

Response	Frequency	Percent	Mean: 3.73
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	3	27.27	
Very Useful	8	72.73	

25. The item map.

Response	Frequency	Percent	Mean: 3.55
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	5	45.45	
Very Useful	6	54.55	

26. The practice activities on making bookmark placements.

Response	Frequency	Percent	Mean: 2.91
Not Useful	0	0.00	
Somewhat Useful	3	27.27	
Useful	6	54.55	
Very Useful	2	18.18	

27. Table-level discussion.

Response	Frequency	Percent	Mean: 3.73
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	3	27.27	
Very Useful	8	72.73	

28. Learning about the other tables' discussions.

Response	Frequency	Percent	Mean: 3.64
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	4	36.36	
Very Useful	7	63.64	

29. Large-group feedback and discussion.

Response	Frequency	Percent	Mean: 3.70
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	3	27.27	
Very Useful	7	63.64	
No Response	1	9.09	

30. The benchmarks.

Response	Frequency	Percent	Mean: 2.91
Not Useful	0	0.00	
Somewhat Useful	3	27.27	
Useful	6	54.55	
Very Useful	2	18.18	

31. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 3.36
Not Useful	0	0.00	
Somewhat Useful	1	9.09	
Useful	5	45.45	
Very Useful	5	45.45	

32. The across-grade presentation of impact data.

Response	Frequency	Percent	Mean: 3.09
Not Useful	0	0.00	
Somewhat Useful	3	27.27	
Useful	4	36.36	
Very Useful	4	36.36	

33. The performance level descriptors (PLDs).

Response	Frequency	Percent	Mean: 3.36
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	7	63.64	
Very Influential	4	36.36	

34. The descriptions of the threshold students.

Response	Frequency	Percent	Mean: 3.64
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	4	36.36	
Very Influential	7	63.64	

35. My perception of the difficulty of the items.

Response	Frequency	Percent	Mean: 2.82
Not Influential	0	0.00	
Somewhat Influential	3	27.27	
Influential	7	63.64	
Very Influential	1	9.09	

36. My experiences with students.

Response	Frequency	Percent	Mean: 3.00
Not Influential	0	0.00	
Somewhat Influential	2	18.18	
Influential	7	63.64	
Very Influential	2	18.18	

37. Discussion at my table.

Response	Frequency	Percent	Mean: 3.36
Not Influential	0	0.00	
Somewhat Influential	1	9.09	
Influential	5	45.45	
Very Influential	5	45.45	

38. Discussion within my group.

Response	Frequency	Percent	Mean: 3.55
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	5	45.45	
Very Influential	6	54.55	

39. The bookmark placements of other participants.

Response	Frequency	Percent	Mean: 2.82
Not Influential	0	0.00	
Somewhat Influential	4	36.36	
Influential	5	45.45	
Very Influential	2	18.18	

40. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 2.55
Not Influential	1	9.09	
Somewhat Influential	4	36.36	
Influential	5	45.45	
Very Influential	1	9.09	

41. The benchmarks.

Response	Frequency	Percent	Mean: 2.55
Not Influential	0	0.00	
Somewhat Influential	6	54.55	
Influential	4	36.36	
Very Influential	1	9.09	

42. My sense of what a student needs to know in order to be Basic.

Response	Frequency	Percent	Mean: 2.91
Not Influential	0	0.00	
Somewhat Influential	2	18.18	
Influential	8	72.73	
Very Influential	1	9.09	

43. My sense of what a student needs to know in order to be Proficient.

Response	Frequency	Percent	Mean: 2.91
Not Influential	0	0.00	
Somewhat Influential	2	18.18	
Influential	8	72.73	
Very Influential	1	9.09	

44. My sense of what a student needs to know in order to be Advanced.

Response	Frequency	Percent	Mean: 2.91
Not Influential	0	0.00	
Somewhat Influential	2	18.18	
Influential	8	72.73	
Very Influential	1	9.09	

45. The general session.

Response	Frequency	Percent	Mean: 2.18
Too Little Time	1	9.09	
About Right	7	63.64	
Too Much Time	3	27.27	

46. Reviewing the test.

Response	Frequency	Percent	Mean: 1.82
Too Little Time	2	18.18	
About Right	9	81.82	
Too Much Time	0	0.00	

47. Reviewing the PLDs.

Response	Frequency	Percent	Mean: 1.73
Too Little Time	3	27.27	
About Right	8	72.73	
Too Much Time	0	0.00	

48. Describing the threshold students.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

49. Reviewing the ordered item booklet (OIB).

Response	Frequency	Percent	Mean: 1.91
Too Little Time	1	9.09	
About Right	10	90.91	
Too Much Time	0	0.00	

50. Reviewing the item map.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

51. Training in the Bookmark Procedure.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

52. Practice activities making bookmark placements.

Response	Frequency	Percent	Mean: 1.91
Too Little Time	2	18.18	
About Right	8	72.73	
Too Much Time	1	9.09	

53. Table discussions after Round 1.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

54. Table discussions after Round 2.

Response	Frequency	Percent	Mean: 2.09
Too Little Time	0	0.00	
About Right	10	90.91	
Too Much Time	1	9.09	

55. Group discussion after Round 2.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

56. Basic cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	11	100.00	
Too High	0	0.00	

57. Proficient cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	11	100.00	
Too High	0	0.00	

58. Advanced cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	11	100.00	
Too High	0	0.00	

59. Basic cut score

Response	Frequency	Percent	Mean: 4.73
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	0	0.00	
Neutral	0	0.00	
Somewhat Comfortable	3	27.27	
Very Comfortable	8	72.73	

60. Proficient cut score

Response	Frequency	Percent	Mean: 4.73
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	0	0.00	
Neutral	0	0.00	
Somewhat Comfortable	3	27.27	
Very Comfortable	8	72.73	

61. Advanced cut score

Response	Frequency	Percent	Mean: 4.82
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	0	0.00	
Neutral	0	0.00	
Somewhat Comfortable	2	18.18	
Very Comfortable	9	81.82	

62. In which group did you work?

Response	Frequency	Percent	Mean: 2.00
Grade 5 Science	0	0.00	
Grade 8 Science	11	100.00	
Physical Science	0	0.00	
Biology	0	0.00	

63. What is your current assignment?

Response	Frequency	Percent	Mean: 1.82
Classroom teacher	8	72.73	
Educator non-teacher	0	0.00	
Higher education	0	0.00	
Other	3	27.27	

64. How many years, in total, have you been teaching?

Response	Frequency	Percent	Mean: 3.73
Fewer than 5 years	0	0.00	
5-10 years	4	36.36	
11-15 years	1	9.09	
16-20 years	1	9.09	
21-25 years	4	36.36	
More than 25 years	1	9.09	

65. What is your gender?

Response	Frequency	Percent	Mean: 1.27
Female	8	72.73	
Male	3	27.27	

66. What is your ethnicity?

Response	Frequency	Percent	Mean: 6.91
American Indian/Alaska Native	0	0.00	
Asian	0	0.00	
Hawaiian or Pacific Islander	0	0.00	
Black	0	0.00	
Hispanic	0	0.00	
Mixed -Two or more races	1	9.09	
Caucasian	10	90.91	

67. What is your highest level of education?

Response	Frequency	Percent	Mean: 5.82
High school diploma	0	0.00	
Associate's degree	0	0.00	
Bachelor's degree	0	0.00	
Bachelor's degree + Hours	1	9.09	
Master's degree	0	0.00	
Master's degree + Hours	10	90.91	
Doctoral degree	0	0.00	

68. Which of these groups do you have experience teaching?

Response	Frequency	Percent	Mean: -
Special education -in a self-contained classroom	0	0.00	
Special education -in a mainstream classroom	7	63.64	
English language learners	4	36.36	
Gifted and talented	5	45.45	
Vocational education	0	0.00	
No Response	2	18.18	

69. What percent of students qualify for free/reduced-price meals at your school?

Response	Frequency	Percent	Mean: 2.64
0-25%	2	18.18	
26-50%	3	27.27	
51-75%	3	27.27	
76% or higher	3	27.27	

Physical Science

1. I understood the purpose of this workshop.

Response	Frequency	Percent	Mean: 4.82
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	18.18	
Strongly Agree	9	81.82	

2. The pre-workshop assignment was helpful for me to engage in the workshop activities.

Response	Frequency	Percent	Mean: 3.78
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	4	36.36	
Agree	3	27.27	
Strongly Agree	2	18.18	
No Response	2	18.18	

3. I understood the content measured by the assessment I reviewed at the standard setting.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	6	54.55	

4. I understood how the assessment was administered.

Response	Frequency	Percent	Mean: 4.45
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	5	45.45	

5. I understood the difference between range PLDs and threshold student descriptors.

Response	Frequency	Percent	Mean: 4.18
Strongly Disagree	1	9.09	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	5	45.45	

6. The PLDs were clear enough for me to describe the threshold students.

Response	Frequency	Percent	Mean: 4.18
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	3	27.27	
Agree	3	27.27	
Strongly Agree	5	45.45	

7. Before Round 1 began, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 4.09
Strongly Disagree	0	0.00	
Disagree	1	9.09	
Neutral	1	9.09	
Agree	5	45.45	
Strongly Agree	4	36.36	

8. I understood how to use the item map.

Response	Frequency	Percent	Mean: 4.45
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	5	45.45	

9. I understood the ordered item booklet (and passage booklet).

Response	Frequency	Percent	Mean: 4.64
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	36.36	
Strongly Agree	7	63.64	

10. I understood how to place my bookmarks.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	3	27.27	
Strongly Agree	7	63.64	

11. I understood the room-level data that was presented between the rounds.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	3	27.27	
Strongly Agree	7	63.64	

12. I understood what the benchmarks were and how I should consider them.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	6	54.55	

13. I understood the impact data that were presented.

Response	Frequency	Percent	Mean: 4.45
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	54.55	
Strongly Agree	5	45.45	

14. By the end of the workshop, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 4.73
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	27.27	
Strongly Agree	8	72.73	

15. The instructions provided in the training materials were clear.

Response	Frequency	Percent	Mean: 4.55
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	3	27.27	
Strongly Agree	7	63.64	

16. The instructions provided during the opening training session were clear.

Response	Frequency	Percent	Mean: 4.64
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	36.36	
Strongly Agree	7	63.64	

17. The instructions provided by the facilitators in my breakout room were clear.

Response	Frequency	Percent	Mean: 4.73
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	27.27	
Strongly Agree	8	72.73	

18. Overall, I believe my opinions were considered and valued by my group.

Response	Frequency	Percent	Mean: 4.73
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	27.27	
Strongly Agree	8	72.73	

19. My group's work was reflected in the presentation of recommendations across grades.

Response	Frequency	Percent	Mean: 4.40
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	9.09	
Agree	4	36.36	
Strongly Agree	5	45.45	
No Response	1	9.09	

20. Overall, I valued the workshop as a professional development experience.

Response	Frequency	Percent	Mean: 4.73
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	27.27	
Strongly Agree	8	72.73	

21. This process will lead to defensible performance standards for the test.

Response	Frequency	Percent	Mean: 4.50
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	45.45	
Strongly Agree	5	45.45	
No Response	1	9.09	

22. Taking and discussing the test before placing a bookmark.

Response	Frequency	Percent	Mean: 3.73
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	3	27.27	
Very Useful	8	72.73	

23. Describing the three categories of threshold students.

Response	Frequency	Percent	Mean: 3.55
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	5	45.45	
Very Useful	6	54.55	

24. Reviewing the ordered item booklet and passage booklet before placing a bookmark.

Response	Frequency	Percent	Mean: 3.73
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	3	27.27	
Very Useful	8	72.73	

25. The item map.

Response	Frequency	Percent	Mean: 3.45
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	6	54.55	
Very Useful	5	45.45	

26. The practice activities on making bookmark placements.

Response	Frequency	Percent	Mean: 2.91
Not Useful	1	9.09	
Somewhat Useful	3	27.27	
Useful	3	27.27	
Very Useful	4	36.36	

27. Table-level discussion.

Response	Frequency	Percent	Mean: 3.82
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	2	18.18	
Very Useful	9	81.82	

28. Learning about the other tables' discussions.

Response	Frequency	Percent	Mean: 3.40
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	6	54.55	
Very Useful	4	36.36	
No Response	1	9.09	

29. Large-group feedback and discussion.

Response	Frequency	Percent	Mean: 3.38
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	5	45.45	
Very Useful	3	27.27	
No Response	3	27.27	

30. The benchmarks.

Response	Frequency	Percent	Mean: 3.55
Not Useful	0	0.00	
Somewhat Useful	1	9.09	
Useful	3	27.27	
Very Useful	7	63.64	

31. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 3.36
Not Useful	0	0.00	
Somewhat Useful	1	9.09	
Useful	5	45.45	
Very Useful	5	45.45	

32. The across-grade presentation of impact data.

Response	Frequency	Percent	Mean: 3.29
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	5	45.45	
Very Useful	2	18.18	
No Response	4	36.36	

33. The performance level descriptors (PLDs).

Response	Frequency	Percent	Mean: 3.73
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	3	27.27	
Very Influential	8	72.73	

34. The descriptions of the threshold students.

Response	Frequency	Percent	Mean: 3.45
Not Influential	1	9.09	
Somewhat Influential	0	0.00	
Influential	3	27.27	
Very Influential	7	63.64	

35. My perception of the difficulty of the items.

Response	Frequency	Percent	Mean: 2.73
Not Influential	1	9.09	
Somewhat Influential	2	18.18	
Influential	7	63.64	
Very Influential	1	9.09	

36. My experiences with students.

Response	Frequency	Percent	Mean: 2.82
Not Influential	1	9.09	
Somewhat Influential	1	9.09	
Influential	8	72.73	
Very Influential	1	9.09	

37. Discussion at my table.

Response	Frequency	Percent	Mean: 3.55
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	5	45.45	
Very Influential	6	54.55	

38. Discussion within my group.

Response	Frequency	Percent	Mean: 3.50
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	5	45.45	
Very Influential	5	45.45	
No Response	1	9.09	

39. The bookmark placements of other participants.

Response	Frequency	Percent	Mean: 2.91
Not Influential	0	0.00	
Somewhat Influential	1	9.09	
Influential	10	90.91	
Very Influential	0	0.00	

40. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 2.73
Not Influential	2	18.18	
Somewhat Influential	2	18.18	
Influential	4	36.36	
Very Influential	3	27.27	

41. The benchmarks.

Response	Frequency	Percent	Mean: 3.09
Not Influential	0	0.00	
Somewhat Influential	3	27.27	
Influential	4	36.36	
Very Influential	4	36.36	

42. My sense of what a student needs to know in order to be Basic.

Response	Frequency	Percent	Mean: 2.91
Not Influential	1	9.09	
Somewhat Influential	2	18.18	
Influential	5	45.45	
Very Influential	3	27.27	

43. My sense of what a student needs to know in order to be Proficient.

Response	Frequency	Percent	Mean: 2.91
Not Influential	1	9.09	
Somewhat Influential	2	18.18	
Influential	5	45.45	
Very Influential	3	27.27	

44. My sense of what a student needs to know in order to be Advanced.

Response	Frequency	Percent	Mean: 2.91
Not Influential	1	9.09	
Somewhat Influential	2	18.18	
Influential	5	45.45	
Very Influential	3	27.27	

45. The general session.

Response	Frequency	Percent	Mean: 2.55
Too Little Time	0	0.00	
About Right	5	45.45	
Too Much Time	6	54.55	

46. Reviewing the test.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

47. Reviewing the PLDs.

Response	Frequency	Percent	Mean: 1.82
Too Little Time	2	18.18	
About Right	9	81.82	
Too Much Time	0	0.00	

48. Describing the threshold students.

Response	Frequency	Percent	Mean: 2.09
Too Little Time	0	0.00	
About Right	10	90.91	
Too Much Time	1	9.09	

49. Reviewing the ordered item booklet (OIB).

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

50. Reviewing the item map.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

51. Training in the Bookmark Procedure.

Response	Frequency	Percent	Mean: 2.18
Too Little Time	0	0.00	
About Right	9	81.82	
Too Much Time	2	18.18	

52. Practice activities making bookmark placements.

Response	Frequency	Percent	Mean: 2.18
Too Little Time	0	0.00	
About Right	9	81.82	
Too Much Time	2	18.18	

53. Table discussions after Round 1.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	11	100.00	
Too Much Time	0	0.00	

54. Table discussions after Round 2.

Response	Frequency	Percent	Mean: 1.82
Too Little Time	2	18.18	
About Right	9	81.82	
Too Much Time	0	0.00	

55. Group discussion after Round 2.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	1	9.09	
About Right	9	81.82	
Too Much Time	1	9.09	

56. Basic cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	11	100.00	
Too High	0	0.00	

57. Proficient cut score

Response	Frequency	Percent	Mean: 2.18
Too Low	1	9.09	
About Right	7	63.64	
Too High	3	27.27	

58. Advanced cut score

Response	Frequency	Percent	Mean: 2.09
Too Low	0	0.00	
About Right	10	90.91	
Too High	1	9.09	

59. Basic cut score

Response	Frequency	Percent	Mean: 4.50
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	0	0.00	
Neutral	2	18.18	
Somewhat Comfortable	1	9.09	
Very Comfortable	7	63.64	
No Response	1	9.09	

60. Proficient cut score

Response	Frequency	Percent	Mean: 4.20
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	2	18.18	
Neutral	0	0.00	
Somewhat Comfortable	2	18.18	
Very Comfortable	6	54.55	
No Response	1	9.09	

61. Advanced cut score

Response	Frequency	Percent	Mean: 4.60
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	0	0.00	
Neutral	0	0.00	
Somewhat Comfortable	4	36.36	
Very Comfortable	6	54.55	
No Response	1	9.09	

62. In which group did you work?

Response	Frequency	Percent	Mean: 3.00
Grade 5 Science	0	0.00	
Grade 8 Science	0	0.00	
Physical Science	11	100.00	
Biology	0	0.00	

63. What is your current assignment?

Response	Frequency	Percent	Mean: 1.55
Classroom teacher	7	63.64	
Educator non-teacher	3	27.27	
Higher education	0	0.00	
Other	1	9.09	

64. How many years, in total, have you been teaching?

Response	Frequency	Percent	Mean: 3.00
Fewer than 5 years	1	9.09	
5-10 years	5	45.45	
11-15 years	1	9.09	
16-20 years	2	18.18	
21-25 years	1	9.09	
More than 25 years	1	9.09	

65. What is your gender?

Response	Frequency	Percent	Mean: 1.36
Female	7	63.64	
Male	4	36.36	

66. What is your ethnicity?

Response	Frequency	Percent	Mean: 7.00
American Indian/Alaska Native	0	0.00	
Asian	0	0.00	
Hawaiian or Pacific Islander	0	0.00	
Black	0	0.00	
Hispanic	0	0.00	
Mixed -Two or more races	0	0.00	
Caucasian	11	100.00	

67. What is your highest level of education?

Response	Frequency	Percent	Mean: 5.73
High school diploma	0	0.00	
Associate's degree	0	0.00	
Bachelor's degree	1	9.09	
Bachelor's degree + Hours	0	0.00	
Master's degree	2	18.18	
Master's degree + Hours	6	54.55	
Doctoral degree	2	18.18	

68. Which of these groups do you have experience teaching?

Response	Frequency	Percent	Mean: -
Special education -in a self-contained classroom	2	18.18	
Special education -in a mainstream classroom	4	36.36	
English language learners	1	9.09	
Gifted and talented	3	27.27	
Vocational education	0	0.00	
No Response	5	45.45	

69. What percent of students qualify for free/reduced-price meals at your school?

Response	Frequency	Percent	Mean: 2.73
0-25%	1	9.09	
26-50%	2	18.18	
51-75%	7	63.64	
76% or higher	1	9.09	

Biology

1. I understood the purpose of this workshop.

Response	Frequency	Percent	Mean: 4.70
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	30.00	
Strongly Agree	7	70.00	

2. The pre-workshop assignment was helpful for me to engage in the workshop activities.

Response	Frequency	Percent	Mean: 3.67
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	5	50.00	
Agree	2	20.00	
Strongly Agree	2	20.00	
No Response	1	10.00	

3. I understood the content measured by the assessment I reviewed at the standard setting.

Response	Frequency	Percent	Mean: 4.80
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	20.00	
Strongly Agree	8	80.00	

4. I understood how the assessment was administered.

Response	Frequency	Percent	Mean: 4.90
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	1	10.00	
Strongly Agree	9	90.00	

5. I understood the difference between range PLDs and threshold student descriptors.

Response	Frequency	Percent	Mean: 4.60
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	40.00	
Strongly Agree	6	60.00	

6. The PLDs were clear enough for me to describe the threshold students.

Response	Frequency	Percent	Mean: 4.60
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	40.00	
Strongly Agree	6	60.00	

7. Before Round 1 began, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 4.40
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	6	60.00	
Strongly Agree	4	40.00	

8. I understood how to use the item map.

Response	Frequency	Percent	Mean: 4.70
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	30.00	
Strongly Agree	7	70.00	

9. I understood the ordered item booklet (and passage booklet).

Response	Frequency	Percent	Mean: 4.80
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	20.00	
Strongly Agree	8	80.00	

10. I understood how to place my bookmarks.

Response	Frequency	Percent	Mean: 4.70
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	30.00	
Strongly Agree	7	70.00	

11. I understood the room-level data that was presented between the rounds.

Response	Frequency	Percent	Mean: 4.90
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	1	10.00	
Strongly Agree	9	90.00	

12. I understood what the benchmarks were and how I should consider them.

Response	Frequency	Percent	Mean: 4.70
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	30.00	
Strongly Agree	7	70.00	

13. I understood the impact data that were presented.

Response	Frequency	Percent	Mean: 4.70
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	3	30.00	
Strongly Agree	7	70.00	

14. By the end of the workshop, I was comfortable with the Bookmark Procedure.

Response	Frequency	Percent	Mean: 5.00
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	0	0.00	
Strongly Agree	10	100.00	

15. The instructions provided in the training materials were clear.

Response	Frequency	Percent	Mean: 4.50
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	50.00	
Strongly Agree	5	50.00	

16. The instructions provided during the opening training session were clear.

Response	Frequency	Percent	Mean: 4.50
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	5	50.00	
Strongly Agree	5	50.00	

17. The instructions provided by the facilitators in my breakout room were clear.

Response	Frequency	Percent	Mean: 4.30
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	7	70.00	
Strongly Agree	3	30.00	

18. Overall, I believe my opinions were considered and valued by my group.

Response	Frequency	Percent	Mean: 4.80
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	20.00	
Strongly Agree	8	80.00	

19. My group's work was reflected in the presentation of recommendations across grades.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	20.00	
Strongly Agree	6	60.00	
No Response	2	20.00	

20. Overall, I valued the workshop as a professional development experience.

Response	Frequency	Percent	Mean: 4.67
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	10.00	
Agree	1	10.00	
Strongly Agree	7	70.00	
No Response	1	10.00	

21. This process will lead to defensible performance standards for the test.

Response	Frequency	Percent	Mean: 4.33
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	10.00	
Agree	4	40.00	
Strongly Agree	4	40.00	
No Response	1	10.00	

22. Taking and discussing the test before placing a bookmark.

Response	Frequency	Percent	Mean: 3.38
Not Useful	0	0.00	
Somewhat Useful	1	10.00	
Useful	3	30.00	
Very Useful	4	40.00	
No Response	2	20.00	

23. Describing the three categories of threshold students.

Response	Frequency	Percent	Mean: 3.63
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	3	30.00	
Very Useful	5	50.00	
No Response	2	20.00	

24. Reviewing the ordered item booklet and passage booklet before placing a bookmark.

Response	Frequency	Percent	Mean: 3.88
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	1	10.00	
Very Useful	7	70.00	
No Response	2	20.00	

25. The item map.

Response	Frequency	Percent	Mean: 3.75
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	2	20.00	
Very Useful	6	60.00	
No Response	2	20.00	

26. The practice activities on making bookmark placements.

Response	Frequency	Percent	Mean: 3.13
Not Useful	0	0.00	
Somewhat Useful	2	20.00	
Useful	3	30.00	
Very Useful	3	30.00	
No Response	2	20.00	

27. Table-level discussion.

Response	Frequency	Percent	Mean: 4.00
Not Useful	0	0.00	
Somewhat Useful	0	0.00	
Useful	0	0.00	
Very Useful	8	80.00	
No Response	2	20.00	

28. Learning about the other tables' discussions.

Response	Frequency	Percent	Mean: 2.88
Not Useful	0	0.00	
Somewhat Useful	2	20.00	
Useful	5	50.00	
Very Useful	1	10.00	
No Response	2	20.00	

29. Large-group feedback and discussion.

Response	Frequency	Percent	Mean: 2.86
Not Useful	0	0.00	
Somewhat Useful	3	30.00	
Useful	2	20.00	
Very Useful	2	20.00	
No Response	3	30.00	

30. The benchmarks.

Response	Frequency	Percent	Mean: 2.71
Not Useful	1	10.00	
Somewhat Useful	2	20.00	
Useful	2	20.00	
Very Useful	2	20.00	
No Response	3	30.00	

31. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 2.75
Not Useful	1	10.00	
Somewhat Useful	2	20.00	
Useful	3	30.00	
Very Useful	2	20.00	
No Response	2	20.00	

32. The across-grade presentation of impact data.

Response	Frequency	Percent	Mean: 2.71
Not Useful	0	0.00	
Somewhat Useful	3	30.00	
Useful	3	30.00	
Very Useful	1	10.00	
No Response	3	30.00	

33. The performance level descriptors (PLDs).

Response	Frequency	Percent	Mean: 4.00
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	0	0.00	
Very Influential	8	80.00	
No Response	2	20.00	

34. The descriptions of the threshold students.

Response	Frequency	Percent	Mean: 3.75
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	2	20.00	
Very Influential	6	60.00	
No Response	2	20.00	

35. My perception of the difficulty of the items.

Response	Frequency	Percent	Mean: 3.25
Not Influential	0	0.00	
Somewhat Influential	1	10.00	
Influential	4	40.00	
Very Influential	3	30.00	
No Response	2	20.00	

36. My experiences with students.

Response	Frequency	Percent	Mean: 3.13
Not Influential	0	0.00	
Somewhat Influential	1	10.00	
Influential	5	50.00	
Very Influential	2	20.00	
No Response	2	20.00	

37. Discussion at my table.

Response	Frequency	Percent	Mean: 3.75
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	2	20.00	
Very Influential	6	60.00	
No Response	2	20.00	

38. Discussion within my group.

Response	Frequency	Percent	Mean: 3.71
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	2	20.00	
Very Influential	5	50.00	
No Response	3	30.00	

39. The bookmark placements of other participants.

Response	Frequency	Percent	Mean: 3.00
Not Influential	0	0.00	
Somewhat Influential	2	20.00	
Influential	4	40.00	
Very Influential	2	20.00	
No Response	2	20.00	

40. The percent of students in each performance level (the impact data).

Response	Frequency	Percent	Mean: 2.75
Not Influential	1	10.00	
Somewhat Influential	2	20.00	
Influential	3	30.00	
Very Influential	2	20.00	
No Response	2	20.00	

41. The benchmarks.

Response	Frequency	Percent	Mean: 3.13
Not Influential	0	0.00	
Somewhat Influential	2	20.00	
Influential	3	30.00	
Very Influential	3	30.00	
No Response	2	20.00	

42. My sense of what a student needs to know in order to be Basic.

Response	Frequency	Percent	Mean: 3.63
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	3	30.00	
Very Influential	5	50.00	
No Response	2	20.00	

43. My sense of what a student needs to know in order to be Proficient.

Response	Frequency	Percent	Mean: 3.63
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	3	30.00	
Very Influential	5	50.00	
No Response	2	20.00	

44. My sense of what a student needs to know in order to be Advanced.

Response	Frequency	Percent	Mean: 3.63
Not Influential	0	0.00	
Somewhat Influential	0	0.00	
Influential	3	30.00	
Very Influential	5	50.00	
No Response	2	20.00	

45. The general session.

Response	Frequency	Percent	Mean: 2.56
Too Little Time	0	0.00	
About Right	4	40.00	
Too Much Time	5	50.00	
No Response	1	10.00	

46. Reviewing the test.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	9	90.00	
Too Much Time	0	0.00	
No Response	1	10.00	

47. Reviewing the PLDs.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	0	0.00	
About Right	9	90.00	
Too Much Time	0	0.00	
No Response	1	10.00	

48. Describing the threshold students.

Response	Frequency	Percent	Mean: 2.00
Too Little Time	1	10.00	
About Right	7	70.00	
Too Much Time	1	10.00	
No Response	1	10.00	

49. Reviewing the ordered item booklet (OIB).

Response	Frequency	Percent	Mean: 2.11
Too Little Time	0	0.00	
About Right	8	80.00	
Too Much Time	1	10.00	
No Response	1	10.00	

50. Reviewing the item map.

Response	Frequency	Percent	Mean: 2.11
Too Little Time	0	0.00	
About Right	8	80.00	
Too Much Time	1	10.00	
No Response	1	10.00	

51. Training in the Bookmark Procedure.

Response	Frequency	Percent	Mean: 2.22
Too Little Time	0	0.00	
About Right	7	70.00	
Too Much Time	2	20.00	
No Response	1	10.00	

52. Practice activities making bookmark placements.

Response	Frequency	Percent	Mean: 2.11
Too Little Time	0	0.00	
About Right	8	80.00	
Too Much Time	1	10.00	
No Response	1	10.00	

53. Table discussions after Round 1.

Response	Frequency	Percent	Mean: 2.11
Too Little Time	0	0.00	
About Right	8	80.00	
Too Much Time	1	10.00	
No Response	1	10.00	

54. Table discussions after Round 2.

Response	Frequency	Percent	Mean: 2.22
Too Little Time	0	0.00	
About Right	7	70.00	
Too Much Time	2	20.00	
No Response	1	10.00	

55. Group discussion after Round 2.

Response	Frequency	Percent	Mean: 2.22
Too Little Time	0	0.00	
About Right	7	70.00	
Too Much Time	2	20.00	
No Response	1	10.00	

56. Basic cut score

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	9	90.00	
Too High	0	0.00	
No Response	1	10.00	

57. Proficient cut score

Response	Frequency	Percent	Mean: 1.89
Too Low	3	30.00	
About Right	4	40.00	
Too High	2	20.00	
No Response	1	10.00	

58. Advanced cut score

Response	Frequency	Percent	Mean: 1.33
Too Low	6	60.00	
About Right	3	30.00	
Too High	0	0.00	
No Response	1	10.00	

59. Basic cut score

Response	Frequency	Percent	Mean: 4.00
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	1	10.00	
Neutral	1	10.00	
Somewhat Comfortable	4	40.00	
Very Comfortable	3	30.00	
No Response	1	10.00	

60. Proficient cut score

Response	Frequency	Percent	Mean: 3.44
Very Uncomfortable	0	0.00	
Somewhat Uncomfortable	2	20.00	
Neutral	1	10.00	
Somewhat Comfortable	6	60.00	
Very Comfortable	0	0.00	
No Response	1	10.00	

61. Advanced cut score

Response	Frequency	Percent	Mean: 2.56
Very Uncomfortable	2	20.00	
Somewhat Uncomfortable	3	30.00	
Neutral	1	10.00	
Somewhat Comfortable	3	30.00	
Very Comfortable	0	0.00	
No Response	1	10.00	

62. In which group did you work?

Response	Frequency	Percent	Mean: 4.00
Grade 5 Science	0	0.00	
Grade 8 Science	0	0.00	
Physical Science	0	0.00	
Biology	9	90.00	
No Response	1	10.00	

63. What is your current assignment?

Response	Frequency	Percent	Mean: 1.56
Classroom teacher	6	60.00	
Educator non-teacher	2	20.00	
Higher education	0	0.00	
Other	1	10.00	
No Response	1	10.00	

64. How many years, in total, have you been teaching?

Response	Frequency	Percent	Mean: 3.44
Fewer than 5 years	0	0.00	
5-10 years	3	30.00	
11-15 years	2	20.00	
16-20 years	1	10.00	
21-25 years	3	30.00	
More than 25 years	0	0.00	
No Response	1	10.00	

65. What is your gender?

Response	Frequency	Percent	Mean: 1.44
Female	5	50.00	
Male	4	40.00	
No Response	1	10.00	

66. What is your ethnicity?

Response	Frequency	Percent	Mean: 7.00
American Indian/Alaska Native	0	0.00	
Asian	0	0.00	
Hawaiian or Pacific Islander	0	0.00	
Black	0	0.00	
Hispanic	0	0.00	
Mixed -Two or more races	0	0.00	
Caucasian	9	90.00	
No Response	1	10.00	

67. What is your highest level of education?

Response	Frequency	Percent	Mean: 5.22
High school diploma	0	0.00	
Associate's degree	0	0.00	
Bachelor's degree	1	10.00	
Bachelor's degree + Hours	1	10.00	
Master's degree	2	20.00	
Master's degree + Hours	5	50.00	
Doctoral degree	0	0.00	
No Response	1	10.00	

68. Which of these groups do you have experience teaching?

Response	Frequency	Percent	Mean: -
Special education -in a self-contained classroom	2	20.00	
Special education -in a mainstream classroom	6	60.00	
English language learners	1	10.00	
Gifted and talented	2	20.00	
Vocational education	0	0.00	
No Response	2	20.00	

69. What percent of students qualify for free/reduced-price meals at your school?

Response	Frequency	Percent	Mean: 2.33
0-25%	1	10.00	
26-50%	4	40.00	
51-75%	4	40.00	
76% or higher	0	0.00	
No Response	1	10.00	

Missouri MAP Science Standard Setting Evaluation: Across-Grade Discussion

The purpose of this evaluation is to help document the process used to review the cut scores across grades during the standard setting for the Missouri MAP science tests. Your opinions and comments are important, as they will provide a basis for judging the quality of this process. *Please do not put your name on this form.* When you have completed the evaluation, please give it to a facilitator. ***Thank you!***

	Part 1: ABOUT THE ACROSS-GRADE DISCUSSION Please *consider the statements below* and mark the level of agreement or disagreement you have with each statement. Please bubble ***only one*** of the five options for each statement.	**Strongly Disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
Overall	1. I understood the purpose of the across-grade discussion	O	O	O	O	O
	2. The facilitator made the across-grade discussion process clear to me.	O	O	O	O	O
	3. I considered the recommendations from my original grade/group during the discussion.	O	O	O	O	O
	4. I considered the content-based expectations for students during the discussion.	O	O	O	O	O
	5. I considered the impact data during the discussion.	O	O	O	O	O
	6. I understood how the impact data were calculated.	O	O	O	O	O
	7. I had enough time to hear about the recommendations made by other groups.	O	O	O	O	O
	8. I had enough time to share the recommendations made by my group.	O	O	O	O	O
	9. Overall, the impact data form a reasonable, explainable pattern across grades.	O	O	O	O	O
	10. Overall, the recommendations reflect appropriately rigorous expectations for students.	O	O	O	O	O
	11. Overall, I believe my opinions were considered and valued by my group.	O	O	O	O	O
	12. My group's work was reflected in the presentation of recommendations across grades.	O	O	O	O	O
	13. This process will lead to defensible performance standards for the test.	O	O	O	O	O
	Please indicate your opinion regarding whether you feel the final, recommended cut scores were **too low, about right, or too high** for each cut score. Please bubble ***only one*** of the three options for each cut score.			**Too Low**	**About Right**	**Too High**
	14. *Basic* cut score for my grade/test			O	O	O
	15. *Proficient* cut score for my grade/test			O	O	O
	16. *Advanced* cut score for my grade/test			O	O	O
	17. *Basic* cut score for the other grade/tests in my content area			O	O	O
	18. *Proficient* cut score for the other grade/tests in my content area			O	O	O
	19. *Advanced* cut score for the other grade/tests in my content area			O	O	O

Part 2: ABOUT YOU

20. In which group did you work?

- O Grade 5 Science
- O Grade 8 Science
- O Physical Science
- O Biology

21. What is your current assignment?

- O Classroom teacher
- O Educator, non-teacher
- O Higher education
- O Other, (please describe):

____

22. How many years, in total, have you been teaching?

- O Fewer than 5 years
- O 5–10 years
- O 11–15 years
- O 16–20 years
- O 21–25 years
- O More than 25 years

23. What is your gender?

- O Female
- O Male

24. What is your ethnicity?

- O American Indian/Alaska Native
- O Asian
- O Hawaiian or Pacific Islander
- O Black
- O Hispanic
- O Mixed (Two or more races)
- O Caucasian

25. What is your highest level of education?

- O High school diploma
- O Associate's degree
- O Bachelor's degree
- O Bachelor's degree + Hours
- O Master's degree
- O Master's degree + Hours
- O Doctoral degree

26. Which of these groups do you have experience teaching?

- O Special education (in a self-contained classroom)
- O Special education (in a mainstream classroom)
- O English language learners
- O Gifted and talented
- O Vocational education

27. What percent of students qualify for free/reduced-price meals at your school?

- O 0 – 25%
- O 26 – 50%
- O 51 – 75%
- O 76% or higher

Part 3: YOUR TURN

In this box, please feel free to add comments about any of your responses, make suggestions to improve future workshops, or tell us what you liked and did not like about this workshop. ***Thank you!***

1. I understood the purpose of the across-grade discussion.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

2. The facilitator made the across-grade discussion process clear to me.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

3. I considered the recommendations from my original grade/group during the discussion.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

4. I considered the content-based expectations for students during the discussion.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

5. I considered the impact data during the discussion.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

6. I understood how the impact data were calculated.

Response	Frequency	Percent	Mean: 4.50
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	4	50.00	
Strongly Agree	4	50.00	

7. I had enough time to hear about the recommendations made by other groups.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

8. I had enough time to share the recommendations made by my group.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

9. Overall, the impact data form a reasonable, explainable pattern across grades.

Response	Frequency	Percent	Mean: 4.25
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	12.50	
Agree	4	50.00	
Strongly Agree	3	37.50	

10. Overall, the recommendations reflect appropriately rigorous expectations for students.

Response	Frequency	Percent	Mean: 4.38
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	12.50	
Agree	3	37.50	
Strongly Agree	4	50.00	

11. Overall, I believe my opinions were considered and valued by my group.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

12. My group's work was reflected in the presentation of recommendations across grades.

Response	Frequency	Percent	Mean: 4.75
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	0	0.00	
Agree	2	25.00	
Strongly Agree	6	75.00	

13. This process will lead to defensible performance standards for the test.

Response	Frequency	Percent	Mean: 4.50
Strongly Disagree	0	0.00	
Disagree	0	0.00	
Neutral	1	12.50	
Agree	2	25.00	
Strongly Agree	5	62.50	

14. Basic cut score for my grade/test

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	8	100.00	
Too High	0	0.00	

15. Proficient cut score for my grade/test

Response	Frequency	Percent	Mean: 2.13
Too Low	0	0.00	
About Right	7	87.50	
Too High	1	12.50	

16. Advanced cut score for my grade/test

Response	Frequency	Percent	Mean: 1.88
Too Low	1	12.50	
About Right	7	87.50	
Too High	0	0.00	

17. Basic cut score for the other grade/tests in my content area

Response	Frequency	Percent	Mean: 2.00
Too Low	0	0.00	
About Right	8	100.00	
Too High	0	0.00	

18. Proficient cut score for the other grade/tests in my content area

Response	Frequency	Percent	Mean: 1.88
Too Low	2	25.00	
About Right	5	62.50	
Too High	1	12.50	

19. Advanced cut score for the other grade/tests in my content area

Response	Frequency	Percent	Mean: 2.13
Too Low	0	0.00	
About Right	7	87.50	
Too High	1	12.50	

20. In which group did you work?

Response	Frequency	Percent	Mean: 2.50
Grade 5 Science	2	25.00	
Grade 8 Science	2	25.00	
Physical Science	2	25.00	
Biology	2	25.00	

21. What is your current assignment?

Response	Frequency	Percent	Mean: 1.13
Classroom teacher	7	87.50	
Educator non-teacher	1	12.50	
Higher education	0	0.00	
Other	0	0.00	

22. How many years, in total, have you been teaching?

Response	Frequency	Percent	Mean: 3.25
Fewer than 5 years	0	0.00	
5-10 years	4	50.00	
11-15 years	1	12.50	
16-20 years	1	12.50	
21-25 years	1	12.50	
More than 25 years	1	12.50	

23. What is your gender?

Response	Frequency	Percent	Mean: 1.38
Female	5	62.50	
Male	3	37.50	

24. What is your ethnicity?

Response	Frequency	Percent	Mean: 7.00
American Indian/Alaska Native	0	0.00	
Asian	0	0.00	
Hawaiian or Pacific Islander	0	0.00	
Black	0	0.00	
Hispanic	0	0.00	
Mixed -Two or more races	0	0.00	
Caucasian	8	100.00	

25. What is your highest level of education?

Response	Frequency	Percent	Mean: 6.00
High school diploma	0	0.00	
Associate's degree	0	0.00	
Bachelor's degree	0	0.00	
Bachelor's degree + Hours	0	0.00	
Master's degree	0	0.00	
Master's degree + Hours	8	100.00	
Doctoral degree	0	0.00	

26. Which of these groups do you have experience teaching?

Response	Frequency	Percent	Mean: -
Special education -in a self-contained classroom	0	0.00	
Special education -in a mainstream classroom	5	62.50	
English language learners	2	25.00	
Gifted and talented	1	12.50	
Vocational education	0	0.00	
No Response	3	37.50	

27. What percent of students qualify for free/reduced-price meals at your school?

Response	Frequency	Percent	Mean: 2.38
0-25%	1	12.50	
26-50%	4	50.00	
51-75%	2	25.00	
76% or higher	1	12.50	

J
Benchmark Recommendations

Recommendations from the Missouri Science Benchmark Panel

On May 2, 2019, the Missouri Department of Elementary and Secondary Instruction (DESE) convened 12 Missouri educators and administrators to recommend benchmarks and associated contextual information for the Missouri Assessment Program (MAP) tests of grade 5 science, grade 8 science, Biology, and Physical Science.

The purpose of the committee was to determine benchmarks and associated contextual information for the upcoming July 2019 standard setting. At the standard setting, Missouri educators will complete the Bookmark Standard Setting Procedure to recommend cut scores for the four science assessments. At the standard setting, participants' decisions will be informed by *benchmarks*: plausible percentages of students who might be classified as *Proficient* and above, based on well-respected, external measures of students' science performance. The associated *contextual information* comprises the facts that standard setting participants should understand to interpret the benchmarks properly.

Prior to the workshop, DESE determined that benchmarks based on the National Assessment of Educational Progress (NAEP) would be appropriate for MAP grades 5 and 8 science. Missouri has a long history of using NAEP-based benchmarks: such benchmarks were last used during the 2018 standard setting for MAP English language arts and mathematics. The MAP grade-level tests of science measure similar content standards to NAEP, and the definitions of *Proficient* are comparable across programs. DESE made no *a priori* determination regarding the benchmarks for the end-of-course (EOC) assessments of Biology and Physical Science.

For grades 5 and 8, the panel examined benchmarks based on the 2015 administration of NAEP, as well as background information on NAEP. This background information included summary test blueprints from NAEP and MAP, MAP performance level descriptors (PLDs), NAEP achievement level descriptors (ALDs), sample items from NAEP, and state and national data from NAEP science in grades 4 and 8. The panel was asked to recommend contextual information for these benchmarks.

For EOC, the panel examined external test data from the 2015 administration of NAEP in grade 12 science, the college and career readiness benchmark for ACT science, and prior year's data from MAP science (based on the older state content standards).The panel was asked to interpret these data to recommend (a) benchmarks for EOC Biology and Physical Science, and (b) the contextual information that should be shared with standard setting participants for these benchmarks.

Contextual Information for NAEP-Based Benchmarks *for* Grades 5 and 8 Science

Standard setting participants should be given contextual information around the NAEP-based benchmarks in three key areas:

1) ***Structure of NAEP.*** Participants should understand that MAP and NAEP measure similar content but have different structures. In particular, participants should understand:
    a. NAEP is given in grades 4 and 8, and MAP tests in grades 5 and 8
    b. Not all Missouri students take NAEP: there is an intricate sampling system for NAEP
    c. NAEP measures and reports science practices separately (i.e., performance task)
    d. Other item types are similar across tests, as illustrated with sample items; although MAP has technology-enhanced items, and NAEP has a performance task

2) ***PLD comparison.*** To help participants compare the performance expectations between NAEP and MAP, a summary paragraph could be added to the *Proficient* descriptor for MAP.
    a. This paragraph should summarize the science content and practices expected.
    b. The parent-friendly descriptors (on the front of the PLD packet) should remain.

3) ***2015 administration.*** Participants should have background information about the timing associated with the 2015 NAEP administration. In particular, participants should understand:
    a. The 2015 NAEP was the first to incorporate updated science standards.
    b. Missouri's science standards were officially adopted in 2016.
    c. In 2015, *only a few* students in Missouri likely received instruction in the new standards.
    d. Classroom instruction in hands-on science practices may have helped Missouri students perform well on the 2015 NAEP, even without explicit instruction in the new standards.
    e. With the new state standards, it is likely (but not guaranteed) that Missouri students would perform better on NAEP in the future.

Benchmark Recommendation and Contextual Information *for* Biology

For Biology, the panel noted that historically around 65% of students were classified as *Proficient* or higher in prior years on the EOC Biology test, and around 40% of students were similarly classified on the tests of grades 5 and 8 science. Accordingly, the panel determined an appropriate benchmark for Biology would fall within these two values.

Some panelists noted that the new science standards emphasize science practices, as embedded in the science content. Because 2019 will be the first year with assessments reflecting this change, Missouri students may not perform as well on the test in 2019 than in previous years. Other panelists noted that the gap between performance on the tests in grades 5, 8, and Biology is currently large, and shrinking this gap may be appropriate given the change in Biology standards.

The panel noted the natural tension between (a) the increased rigor associated with the new science standards, and (b) the desire for many students to be classified as *Proficient* for the purposes of local accountability. Some panelists noted that the benchmarks for grades 5 and 8 were based on external measures of student data, and that the EOC benchmarks should also be based on external data points.

The panel examined NAEP data from the 2015 administration of the national grade 12 science test, prior year's MAP data, ACT data, and the content measured by each. The panel noted that it was unlikely that Missouri students would perform as well on Biology as observed in 2017 (given the shift in content standards). The panel also noted that the benchmarks for grades 5 and 8 were approximately 40% *Proficient* and above.

The panel did not feel it could recommend a single-point benchmark for the Biology standard setting. Accordingly, the panel recommended a benchmark range from 45% to 50%. (The 45% bound represents a few points above the benchmark for grade 8, acknowledging that students have historically performed better on EOC Biology than on the grade 8 science test; and the 5% width of this band acknowledges that no single point-estimate is warranted.)

Contextual information on this benchmark includes the following:

- The benchmarks for grades 5 and 8 and for EOC are qualitatively different, and the benchmarks for EOC should be interpreted differently than the benchmarks for grades 5 and 8.
- Participants should understand the evolution of the MAP tests. This may be the first year that some teachers used the new state standards.
- There have been critical shifts in the standards and learning expectations starting this year for Biology. This is a large shift in impact data, but it is warranted in terms of the content.

Benchmark Recommendation and Contextual Information *for* Physical Science

The panel noted that many fewer students take the Physical Science test than take the Biology test, and that the performance of examinees was much worse on Physical Science than on Biology. Panelists noted that in many systems, the students who take the Physical Science course have qualitatively different scholastic trajectories than do students who take Biology. The panel did not expect that the percentage of students classified as *Proficient* and above would be as high for Physical Science as for Biology.

The panel looked at the percentage of students classified as *Proficient* and above on grade 8 NAEP, grade 12 NAEP, and the previously-recommended benchmarks for grade 5 and 8 science and Biology.

As for Biology, the panel chose to recommend a benchmark range instead of a point-recommendation. The panel recommended a range of 25% to 35%. (The 25% bound represents a percentage just below the observed impact data from 2016 and 2017. The 10% width of the band recognizes that fewer external data points are available for Physical Science than for other tests.)

Contextual information on this benchmark includes the following:

- As a group, the students who take Physical Science are different than those who take Biology.
- Historically, fewer students have been classified as *Proficient* or above on Physical Science than on Biology.
- …plus the same contextual information for *Biology*.

Appendix D: Item Statistics

Table D.1. Item Statistics—English I

UIN	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0001781	0.60	0.24	0.20
MO0001782	0.62	0.34	0.41
MO0001783	0.50	0.15	0.39
MO0001812	0.85	0.41	0.41
MO0001818	0.32	0.37	0.00
MO0001870	0.71	0.51	0.00
MO0001870_S1	0.62	0.46	0.00
MO0001870_S2	0.68	0.49	0.00
MO0001870_S3	0.96	0.36	0.00
MO0001873	0.68	0.56	0.00
MO0001873_S1	0.59	0.53	0.00
MO0001873_S2	0.64	0.56	0.00
MO0001873_S3	0.94	0.40	0.00
MO0007207	0.60	0.19	0.41
MO0007208	0.79	0.46	0.43
MO0007308	0.46	0.22	0.22
MO0007677	0.40	0.18	0.00
MO0007794	0.47	0.12	0.25
MO0007858	0.52	0.32	0.29
MO0007902	0.56	0.45	0.23
MO0008149	0.50	0.35	0.27
MO0008157	0.66	0.42	0.34
MO0008165	0.44	0.35	0.00
MO0008223	0.62	0.28	0.27
MO0008224	0.53	0.40	0.30
MO0008256	0.52	0.23	0.30
MO0008270	0.47	0.35	0.00
MO0008285	0.41	0.35	0.32
MO0008334	0.51	0.36	0.27
MO0008336	0.35	0.39	0.30
MO0008337	0.53	0.31	0.30
MO0008339	0.55	0.44	0.00
MO0008349	0.61	0.38	0.32
MO0008423	0.52	0.32	0.22
MO0008458	0.66	0.37	0.00
MO0008726	0.49	0.33	0.00
MO0008735	0.44	0.24	0.20
MO0008772	0.57	0.34	0.20
MO0008779	0.49	0.28	0.00
MO0008780	0.54	0.23	0.20
MO0008792	0.46	0.11	0.14
MO0008804	0.78	0.42	0.14
MO0008908	0.65	0.23	0.20
MO0017550	0.83	0.39	0.12
MO0017581	0.74	0.44	0.12
MO0017631	0.46	0.30	0.18
MO0017637	0.70	0.44	0.18
MO0017638	0.57	0.47	0.18
MO0018251	0.64	0.30	0.14
MO0018258	0.52	0.20	0.14
MO0018266	0.76	0.46	0.14

UIN	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0018281	0.72	0.43	0.14
MO0018534	0.63	0.37	0.14
MO0044055	0.54	0.29	0.39
MO0044056	0.58	0.48	0.00
MO0044057	0.76	0.39	0.43
MO0044058	0.76	0.38	0.41
MO0044060	0.47	0.26	0.41
MO0044061	0.55	0.49	0.00
MO0044069	0.76	0.52	0.38
MO0044075	0.64	0.31	0.38
MO0044077	0.79	0.33	0.43
MOE1161	0.60	0.21	0.60
MOE11610	0.67	0.38	0.60
MOE11614	0.72	0.41	0.24
MOE116142	0.65	0.31	0.10
MOE116145	0.03	0.10	0.00
MOE116148	0.50	0.23	0.10
MOE116149	0.61	0.36	0.29
MOE116152	0.68	0.30	0.28
MOE116161	0.56	0.35	0.24
MOE11617	0.50	0.39	0.28
MOE1162	0.25	0.18	0.60
MOE116214	0.76	0.42	0.00
MOE116223	0.14	0.19	0.00
MOE116225	0.59	0.42	0.30
MOE116228	0.53	0.29	0.30
MOE116298	0.27	0.28	0.00
MOE116299	0.77	0.37	0.23
MOE116364	0.64	0.52	0.00
MOE116365	0.18	0.21	0.00
MOE116366	0.56	0.15	0.10
MOE116428	0.78	0.43	0.30
MOE116432	0.78	0.43	0.30
MOE116441	0.58	0.35	0.23
MOE116449	0.53	0.33	0.70
MOE11675	0.56	0.27	0.24
MOE116778	0.61	0.49	0.41
MOE11678	0.55	0.21	0.26
MOE116783	0.52	0.23	0.41
MOE116785	0.42	0.30	0.62
MOE116790	0.59	0.30	0.60
MOE11680	0.51	0.25	0.24
MOE11688	0.65	0.44	0.30
MOE11689	0.57	0.26	0.28

Table D.2. Item Statistics—English II

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0001817	0.69	0.54	0.00
MO0001817_S1	0.60	0.51	0.00
MO0001817_S2	0.65	0.54	0.00
MO0001817_S3	0.96	0.35	0.00
MO0001825	0.47	0.26	0.11
MO0001828	0.78	0.56	0.09
MO0001831	0.71	0.43	0.11
MO0001834	0.70	0.42	0.09
MO0001843	0.70	0.59	0.00
MO0001843_S1	0.60	0.57	0.00
MO0001843_S2	0.67	0.55	0.00
MO0001843_S3	0.96	0.35	0.00
MO0007288	0.42	0.07	0.04
MO0007296	0.48	0.31	0.02
MO0007627	0.52	0.38	0.09
MO0007733	0.67	0.20	0.02
MO0007842	0.41	0.12	0.06
MO0007879	0.66	0.37	0.01
MO0007941	0.73	0.39	0.01
MO0008025	0.36	0.32	0.04
MO0008042	0.78	0.18	0.01
MO0008070	0.78	0.39	0.00
MO0008076	0.61	0.25	0.04
MO0008121	0.60	0.39	0.08
MO0008146	0.64	0.42	0.00
MO0008162	0.39	0.36	0.00
MO0008177	0.54	0.44	0.03
MO0008185	0.51	0.35	0.00
MO0008207	0.36	0.17	0.04
MO0008211	0.41	0.24	0.09
MO0008236	0.51	0.41	0.09
MO0008352	0.54	0.47	0.13
MO0008485	0.39	0.30	0.02
MO0008495	0.61	0.22	0.02
MO0008716	0.58	0.40	0.07
MO0008742	0.61	0.42	0.07
MO0008757	0.74	0.46	0.07
MO0008765	0.67	0.39	0.03
MO0008791	0.61	0.39	0.02
MO0008813	0.27	0.11	0.04
MO0008827	0.40	0.14	0.04
MO0008859	0.76	0.29	0.02
MO0008867	0.51	0.22	0.07
MO0008877	0.45	0.37	0.07
MO0008956	0.46	0.34	0.14
MO0008971	0.30	0.14	0.03
MO0009087	0.83	0.27	0.03
MO0009090	0.24	0.18	0.05
MO0009103	0.32	0.34	0.00
MO0044085	0.81	0.40	0.00
MO0044091	0.71	0.44	0.07
MO0044092	0.60	0.52	0.14
MO0044101	0.55	0.28	0.06

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MOE116160	0.58	0.22	0.02
MOE116291	0.63	0.45	0.02
MOE116294	0.33	0.34	0.00
MOE116295	0.92	0.39	0.02
MOE116296	0.30	0.34	0.00
MOE116300	0.53	0.31	0.10
MOE116303	0.16	0.25	0.00
MOE116354	0.28	0.35	0.00
MOE116355	0.58	0.45	0.10
MOE116358	0.55	0.35	0.09
MOE2161	0.42	0.28	0.08
MOE21612	0.27	0.30	0.00
MOE21619	0.62	0.29	0.00
MOE21620	0.70	0.41	0.06
MOE21621	0.37	0.25	0.04
MOE216215	0.36	0.36	0.09
MOE216217	0.49	0.31	0.09
MOE216218	0.42	0.21	0.00
MOE21622	0.50	0.27	0.00
MOE216220	0.41	0.26	0.07
MOE216224	0.53	0.08	0.07
MOE216237	0.49	0.35	0.05
MOE21624	0.64	0.31	0.02
MOE216303	0.52	0.26	0.04
MOE2165	0.64	0.30	0.07
MOE2166	0.44	0.22	0.08
MOE216709	0.80	0.47	0.07
MOE21675	0.72	0.30	0.04
MOE216775	0.65	0.26	0.06
MOE216776	0.74	0.46	0.04
MOE216779	0.72	0.45	0.04
MOE21678	0.62	0.39	0.05
MOE216788	0.64	0.48	0.10
MOE216791	0.35	0.40	0.00
MOE216793	0.47	0.36	0.00
MOE21683	0.66	0.40	0.04

Table D.3. Item Statistics—Algebra I

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0003373	0.34	0.39	0.06
MO0003682	0.58	0.41	0.05
MO0003888	0.27	0.26	0.11
MO0007286	0.08	0.38	0.00
MO0007747	0.38	0.24	0.12
MO0007829	0.85	0.20	0.00
MO0007870	0.51	0.58	0.00
MO0007928	0.38	0.22	0.29
MO0007929	0.39	0.25	0.23
MO0008117	0.36	0.32	0.12
MO0008160	0.45	0.33	0.23
MO0008164	0.52	0.53	0.31
MO0008175	0.68	0.50	0.17
MO0008200	0.26	0.27	0.31
MO0008228	0.42	0.29	0.20
MO0008294	0.48	0.48	0.00
MO0008313	0.67	0.50	0.00
MO0008325	0.34	0.53	0.00
MO0008357	0.15	0.49	0.00
MO0008366	0.52	0.27	0.32
MO0008367	0.68	0.44	0.23
MO0008370	0.51	0.48	0.00
MO0008379	0.32	0.46	0.00
MO0008405	0.38	0.38	0.18
MO0008411	0.36	0.49	0.29
MO0008414	0.48	0.33	0.07
MO0008437	0.67	0.56	0.00
MO0008446	0.27	0.44	0.00
MO0008463	0.46	0.20	0.18
MO0008476	0.82	0.27	0.02
MO0008481	0.40	0.32	0.32
MO0008491	0.17	0.46	0.00
MO0008494	0.52	0.30	0.10
MO0008519	0.14	0.30	0.00
MO0008715	0.46	0.31	0.28
MO0008731	0.58	0.62	0.10
MO0008732	0.31	0.62	0.00
MO0008734	0.44	0.62	0.00
MO0008746	0.54	0.41	0.17
MO0008747	0.45	0.54	0.00
MO0008749	0.66	0.23	0.00
MO0008762	0.13	0.54	0.00
MO0008766	0.30	0.26	0.00
MO0008767	0.53	0.37	0.06
MO0008774	0.46	0.30	0.32
MO0008794	0.40	0.50	0.00
MO0008811	0.49	0.42	0.30
MO0008893	0.62	0.33	0.28
MO0008907	0.76	0.18	0.01
MO0008925	0.37	0.36	0.00
MO0008961	0.46	0.35	0.27
MO0008965	0.44	0.48	0.00
MO0008969	0.44	0.34	0.23

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0008998	0.39	0.64	0.00
MO0009019	0.52	0.57	0.07
MO0009069	0.57	0.49	0.34
MO0009099	0.33	0.57	0.34
MO0009146	0.23	0.31	0.14
MO0009204	0.78	0.38	0.02
MO0009213	0.45	0.36	0.31
MO0009217	0.66	0.41	0.10
MO0009230	0.33	0.28	0.27
MO0009238	0.42	0.23	0.17
MO0009242	0.73	0.41	0.22
MOA11610	0.42	0.39	1.20
MOA11612	0.35	0.20	0.10
MOA116148	0.28	0.20	0.10
MOA116150	0.84	0.35	0.03
MOA116153	0.42	0.26	0.12
MOA116156	0.57	0.45	0.08
MOA116158	0.43	0.29	0.10
MOA116159	0.42	0.37	0.32
MOA116225	0.67	0.30	0.04
MOA116290	0.48	0.26	0.16
MOA116296	0.68	0.46	0.04
MOA116353	0.36	0.52	0.00
MOA116425	0.36	0.51	0.00
MOA116427	0.68	0.37	0.04
MOA116438	0.30	0.16	0.04
MOA116496	0.53	0.23	0.10
MOA116498	0.21	0.29	0.14
MOA1165	0.60	0.29	0.02
MOA116501	0.47	0.52	0.23
MOA116502	0.26	0.23	0.00
MOA1166	0.57	0.36	0.08
MOA116635	0.34	0.47	0.00
MOA116637	0.82	0.35	0.08
MOA116646	0.53	0.33	0.14
MOA116710	0.54	0.46	1.18
MOA116718	0.27	0.46	0.00
MOA11675	0.39	0.34	0.08
MOA11676	0.52	0.21	0.12
MOA116798	0.32	0.50	0.14
MOA11684	0.25	0.46	0.00
MOA11685	0.35	0.27	0.10

Table D.4. Item Statistics for Algebra II

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0003376	0.25	0.45	0.00
MO0003378	0.66	0.49	0.00
MO0007275	0.80	0.48	0.00
MO0007276	0.56	0.52	0.00
MO0007277	0.51	0.44	0.00
MO0007311	0.42	0.22	0.04
MO0007593	0.44	0.33	0.10
MO0007596	0.43	0.52	0.00
MO0007600	0.35	0.34	0.04
MO0007603	0.34	0.46	0.00
MO0007625	0.59	0.36	0.04
MO0007642	0.69	0.39	0.07
MO0007652	0.41	0.28	0.08
MO0007659	0.71	0.35	0.01
MO0007666	0.47	0.22	0.09
MO0007678	0.76	0.49	0.00
MO0007730	0.54	0.41	0.03
MO0007741	0.53	0.49	0.07
MO0007746	0.38	0.23	0.04
MO0007754	0.51	0.39	0.05
MO0007755	0.64	0.38	0.03
MO0007760	0.70	0.37	0.05
MO0007770	0.34	0.34	0.00
MO0007785	0.43	0.31	0.08
MO0007791	0.43	0.40	0.07
MO0007793	0.39	0.49	0.00
MO0007823	0.72	0.44	0.03
MO0007826	0.59	0.19	0.01
MO0007831	0.66	0.42	0.03
MO0007841	0.30	0.31	0.03
MO0007843	0.33	0.27	0.00
MO0007860	0.59	0.36	0.06
MO0007865	0.37	0.58	0.00
MO0007868	0.49	0.44	0.00
MO0007878	0.24	0.45	0.07
MO0007893	0.40	0.17	0.00
MO0007915	0.41	0.33	0.00
MO0007939	0.50	0.43	0.06
MO0007998	0.52	0.42	0.05
MO0008003	0.67	0.42	0.03
MO0008066	0.71	0.54	0.00
MO0008067	0.35	0.60	0.00
MO0008143	0.48	0.53	0.00
MO0008216	0.36	0.52	0.00
MO0008260	0.70	0.53	0.01
MO0008261	0.49	0.32	0.07
MO0008281	0.33	0.39	0.00
MO0008298	0.37	0.50	0.00
MO0008305	0.40	0.43	0.07
MO0008314	0.72	0.57	0.00

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0008348	0.63	0.26	0.05
MO0008372	0.64	0.37	0.06
MO0008380	0.48	0.11	0.03
MO0020031	0.52	0.38	0.05
MO0020109	0.47	0.57	0.00
MO0020625	0.34	0.49	0.00
MOA2161	0.50	0.31	0.07
MOA216105	0.50	0.42	0.28
MOA216126	0.49	0.46	0.00
MOA216154	0.63	0.51	0.00
MOA21619	0.67	0.45	0.00
MOA21620	0.33	0.42	0.00
MOA216218	0.51	0.46	0.00
MOA216224	0.95	0.16	0.01
MOA216231	0.42	0.51	0.00
MOA216353	0.29	0.48	0.00
MOA216355	0.67	0.45	0.00
MOA216367	0.58	0.39	0.00
MOA216371	0.20	0.46	0.00
MOA216375	0.32	0.42	0.05
MOA216376	0.43	0.50	0.00
MOA216426	0.36	0.23	0.03
MOA216427	0.66	0.36	0.00
MOA216432	0.65	0.57	0.00
MOA216439	0.61	0.43	0.00
MOA216441	0.59	0.45	0.07
MOA216445	0.73	0.23	0.00
MOA216446	0.74	0.43	0.07
MOA216492	0.20	0.47	0.00
MOA216496	0.36	0.06	0.05
MOA216497	0.51	0.44	0.00
MOA216498	0.22	0.49	0.00
MOA216501	0.26	0.17	0.10
MOA216518	0.44	0.15	0.01
MOA216524	0.44	0.31	0.00
MOA21671	0.32	0.33	0.00
MOA21677	0.67	0.32	0.21
MOA21678	0.47	0.26	0.07
MOA21682	0.52	0.58	0.00
MOA2169	0.57	0.49	0.07
MOA21690	0.48	0.35	0.00
MOA21691	0.83	0.27	0.05
MOA21692	0.27	0.22	0.00
MOA21699	0.40	0.29	0.14

Table D.5. Item Statistics—Geometry

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0007136	0.27	0.31	0.00
MO0007142	0.42	0.28	0.00
MO0007197	0.66	0.29	0.00
MO0007227	0.30	0.36	0.00
MO0007248	0.42	0.52	0.00
MO0007598	0.46	0.17	0.00
MO0007628	0.45	0.53	0.00
MO0007645	0.46	0.19	0.00
MO0007695	0.66	0.35	0.00
MO0007708	0.37	0.37	0.00
MO0007715	0.31	0.42	0.00
MO0007716	0.75	0.31	0.00
MO0007736	0.37	0.26	0.00
MO0007745	0.38	0.15	0.03
MO0007757	0.50	0.55	0.00
MO0007784	0.44	0.20	0.00
MO0007911	0.59	0.47	0.00
MO0007913	0.62	0.47	0.00
MO0007937	0.51	0.31	0.00
MO0007943	0.39	0.41	0.00
MO0007944	0.71	0.46	0.00
MO0007946	0.74	0.41	0.00
MO0007969	0.47	0.42	0.00
MO0007975	0.35	0.28	0.00
MO0007978	0.30	0.51	0.00
MO0007994	0.50	0.40	0.00
MO0008010	0.70	0.36	0.00
MO0008022	0.25	0.59	0.00
MO0008029	0.21	0.58	0.00
MO0008046	0.12	0.43	0.00
MO0008060	0.57	0.38	0.00
MO0008072	0.26	0.21	0.00
MO0008082	0.23	0.21	0.00
MO0008098	0.60	0.42	0.00
MO0008180	0.47	0.31	0.00
MO0008206	0.60	0.01	0.00
MO0008291	0.37	0.41	0.00
MO0008311	0.50	0.10	0.00
MO0008312	0.30	0.57	0.06
MO0008320	0.52	0.49	0.00
MO0008328	0.39	0.38	0.00
MO0008427	0.32	0.35	0.00
MO0008754	0.61	0.43	0.00
MO0008784	0.37	0.41	0.00
MO0008829	0.28	0.46	0.00
MO0008861	0.12	0.30	0.00
MO0008865	0.61	0.40	0.00
MO0008871	0.30	0.50	0.00
MO0008938	0.14	0.31	0.00
MO0008939	0.42	0.19	0.00
MO0008946	0.58	0.43	0.00
MO0008979	0.84	0.33	0.00
MO0008981	0.28	0.33	0.00

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0009032	0.64	0.43	0.00
MO0009046	0.31	0.30	0.00
MO0009056	0.29	0.28	0.07
MO0009074	0.53	0.26	0.00
MO0009079	0.44	0.47	0.07
MO0009125	0.56	0.63	0.00
MOG161	0.54	0.34	0.00
MOG1611	0.56	0.38	0.00
MOG1614	0.70	0.42	0.00
MOG16141	0.49	0.30	0.00
MOG16142	0.30	0.36	0.00
MOG16145	0.41	0.39	0.00
MOG16146	0.32	0.43	0.00
MOG16151	0.82	0.34	0.00
MOG16153	0.67	0.37	0.00
MOG1617	0.40	0.17	0.00
MOG16212	0.45	0.53	0.00
MOG16221	0.52	0.37	0.00
MOG16226	0.25	0.16	0.00
MOG16357	0.69	0.42	0.00
MOG16358	0.23	0.34	0.00
MOG16360	0.47	0.31	0.00
MOG16362	0.27	0.07	0.00
MOG16364	0.30	0.48	0.00
MOG16365	0.70	0.28	0.00
MOG16368	0.25	0.31	0.00
MOG16426	0.58	0.37	0.00
MOG16431	0.50	0.16	0.00
MOG16433	0.29	0.46	0.00
MOG16434	0.42	0.36	0.00
MOG16439	0.75	0.30	0.06
MOG16498	0.81	0.26	0.00
MOG16508	0.47	0.46	0.00
MOG166	0.39	0.27	0.32
MOG16775	0.48	0.15	0.07
MOG16777	0.26	0.45	0.00
MOG16778	0.41	0.44	0.00
MOG16783	0.27	0.44	0.00
MOG16788	0.47	0.54	0.00
MOG16789	0.46	0.54	0.00
MOG16791	0.41	0.50	0.00
MOG16801	0.44	0.41	0.00
MOG16805	0.46	0.23	0.00
MOG16810	0.37	0.40	0.00
MOG16811	0.29	0.37	0.00

Table D.6. Item Statistics—Biology

UIN	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0007089	0.51	0.44	0.08
MO0007103	0.67	0.51	0.02
MO0007106	0.74	0.47	0.12
MO0007109	0.38	0.49	0.00
MO0007112	0.34	0.33	0.05
MO0007117	0.64	0.26	0.02
MO0007126	0.48	0.45	0.08
MO0007148	0.49	0.34	0.10
MO0007150	0.79	0.53	0.00
MO0007157	0.56	0.50	0.05
MO0007160	0.53	0.47	0.11
MO0007177	0.80	0.51	0.00
MO0007189	0.62	0.52	0.10
MO0007217	0.63	0.48	0.12
MO0007219	0.37	0.44	0.00
MO0007232	0.29	0.46	0.00
MO0007254	0.50	0.60	0.00
MO0007305	0.62	0.51	0.00
MO0007612	0.49	0.36	0.12
MO0007660	0.78	0.43	0.03
MO0007675	0.62	0.50	0.04
MO0007709	0.34	0.13	0.09
MO0007744	0.40	0.22	0.13
MO0007749	0.35	0.13	0.06
MO0007763	0.34	0.23	0.00
MO0007873	0.57	0.41	0.05
MO0007935	0.57	0.47	0.05
MO0007948	0.82	0.40	0.02
MO0007965	0.67	0.48	0.00
MO0007968	0.70	0.41	0.10
MO0008203	0.63	0.35	0.04
MO0008231	0.71	0.44	0.04
MO0008240	0.77	0.49	0.06
MO0008247	0.53	0.34	0.13
MO0008273	0.53	0.46	0.10
MO0008393	0.66	0.50	0.05
MO0008484	0.50	0.38	0.00
MO0008710	0.40	0.38	0.08
MO0008797	0.43	0.37	0.09
MO0008840	0.76	0.47	0.10
MO0008851	0.44	0.15	0.10
MO0008855	0.77	0.28	0.07
MO0008882	0.70	0.43	0.07
MO0008904	0.76	0.46	0.00
MO0008910	0.75	0.48	0.06
MO0008915	0.56	0.46	0.06
MO0008957	0.71	0.53	0.07
MO0008963	0.57	0.30	0.00
MO0008970	0.47	0.62	0.00
MO0008972	0.65	0.39	0.07
MO0008984	0.54	0.18	0.09
MO0009018	0.64	0.44	0.03
MO0009020	0.78	0.42	0.06

UIN	*p*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0009042	0.90	0.33	0.04
MO0009096	0.88	0.47	0.06
MO0009119	0.56	0.53	0.12
MO0009144	0.84	0.58	0.05
MO0009202	0.61	0.28	0.07
MO0009214	0.54	0.23	0.11
MO0009226	0.52	0.47	0.11
MO0013608	0.62	0.41	0.03
MO0013657	0.52	0.52	0.00
MO0014514	0.74	0.35	0.00
MO0015259	0.73	0.33	0.12
MO0015260	0.35	0.14	0.08
MO0015262	0.75	0.45	0.11
MO0015488	0.53	0.47	0.12
MO0015580	0.84	0.40	0.04
MO0030658	0.59	0.57	0.08
MO0030876	0.74	0.49	0.00
MO0044419	0.80	0.39	0.04
MO0045927	0.58	0.43	0.07

Table D.7. Item Statistics—Physical Science

UIN	*P*-Value/Mean	Corrected Point-Biserial Correlation	Omit Rate (%)
MO0007121	0.18	0.46	0.00
MO0007146	0.67	0.41	0.04
MO0007170	0.56	0.35	0.00
MO0007179	0.58	0.42	0.04
MO0007185	0.56	0.42	0.00
MO0007615	0.41	0.22	0.00
MO0007626	0.62	0.36	0.00
MO0007638	0.50	0.51	0.00
MO0007639	0.42	0.30	0.00
MO0007664	0.64	0.47	0.04
MO0007698	0.45	0.34	0.00
MO0007713	0.90	0.25	0.00
MO0007728	0.46	0.37	0.00
MO0007734	0.66	0.38	0.00
MO0007778	0.65	0.23	0.00
MO0007804	0.47	0.27	0.00
MO0007820	0.41	0.48	0.00
MO0007897	0.70	0.38	0.04
MO0008106	0.44	0.36	0.00
MO0008212	0.76	0.40	0.00
MO0008284	0.57	0.18	0.04
MO0008299	0.57	0.46	0.00
MO0008385	0.49	0.51	0.00
MO0008398	0.61	0.40	0.00
MO0008421	0.78	0.38	0.00
MO0008483	0.74	0.27	0.04
MO0008790	0.49	0.31	0.00
MO0008880	0.62	0.08	0.00
MO0008883	0.47	0.10	0.00
MO0008902	0.51	0.17	0.00
MO0008959	0.62	0.40	0.04
MO0008987	0.65	0.38	0.00
MO0009068	0.40	0.29	0.00
MO0009071	0.46	0.36	0.00
MO0009095	0.45	0.16	0.04
MO0009102	0.84	0.35	0.00
MO0009210	0.78	0.45	0.00
MO0009224	0.57	0.40	0.00
MO0022372	0.42	0.44	0.00
MO0044943	0.58	0.34	0.00

Appendix E: Raw-to-Scale Score (RSS) Conversions

Table E.1. English I Forms

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
0	325	25	1	325	25	1	325	25	1
1	330	15	1	325	25	1	325	25	1
2	341	11	1	325	25	1	325	25	1
3	347	9	1	329	15	1	327	15	1
4	352	8	1	341	11	1	338	11	1
5	356	7	1	348	10	1	345	9	1
6	360	7	1	353	8	1	351	8	1
7	363	7	1	357	8	1	355	8	1
8	365	6	1	361	7	1	359	7	1
9	368	6	1	364	7	1	362	7	1
10	370	6	1	367	7	1	365	6	1
11	372	6	1	370	6	1	368	6	1
12	375	6	1	373	6	1	370	6	1
13	377	5	1	375	6	1	372	6	1
14	379	5	1	377	6	1	375	6	1
15	381	5	1	379	6	1	377	6	1
16	382	5	1	382	5	1	379	5	1
17	384	5	2	383	5	1	381	5	1
18	386	5	2	385	5	2	382	5	1
19	388	5	2	387	5	2	384	5	2
20	389	5	2	389	5	2	386	5	2
21	391	5	2	391	5	2	388	5	2
22	393	5	2	392	5	2	389	5	2
23	394	5	2	394	5	2	391	5	2
24	396	5	2	395	5	2	393	5	2
25	397	5	2	397	5	2	394	5	2
26	399	5	2	399	5	2	396	5	2
27	401	5	3	400	5	3	397	5	2
28	402	5	3	402	5	3	399	5	2
29	404	5	3	403	5	3	401	5	3
30	406	5	3	405	5	3	402	5	3
31	407	5	3	407	5	3	404	5	3
32	409	5	3	408	5	3	406	5	3
33	411	5	3	410	5	3	407	5	3
34	413	5	3	412	5	3	409	5	3
35	415	5	4	414	5	3	411	5	3
36	417	6	4	416	6	4	413	5	3

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
37	419	6	4	418	6	4	415	6	4
38	421	6	4	420	6	4	417	6	4
39	423	6	4	423	6	4	419	6	4
40	426	6	4	425	6	4	422	6	4
41	429	7	4	428	7	4	424	6	4
42	432	7	4	431	7	4	427	7	4
43	435	7	4	435	8	4	431	7	4
44	439	8	4	439	8	4	435	8	4
45	443	8	4	444	9	4	439	9	4
46	448	9	4	449	9	4	445	10	4
47	455	10	4	456	10	4	452	11	4
48	463	12	4	464	12	4	461	13	4
49	476	16	4	477	16	4	475	16	4
50	495	25	4	497	25	4	495	25	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate the achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Table E.2. English II Forms

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
0	325	25	1	325	25	1	325	25	1
1	335	15	1	325	25	1	330	15	1
2	345	10	1	325	25	1	325	25	1
3	351	9	1	328	15	1	330	15	1
4	355	8	1	340	11	1	342	11	1
5	359	7	1	347	9	1	349	9	1
6	362	6	1	352	8	1	354	8	1
7	365	6	1	357	8	1	358	8	1
8	367	6	1	361	7	1	362	7	1
9	370	6	1	364	7	1	365	7	1
10	372	6	1	367	6	1	368	6	1
11	374	5	1	370	6	1	371	6	1
12	376	5	1	373	6	1	373	6	1
13	378	5	1	375	6	1	375	6	1
14	379	5	1	377	6	1	377	5	1
15	381	5	1	379	5	1	379	5	1
16	383	5	1	381	5	1	381	5	1
17	384	5	2	383	5	1	383	5	1
18	386	5	2	385	5	2	385	5	2

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
19	388	5	2	387	5	2	386	5	2
20	389	5	2	389	5	2	388	5	2
21	391	5	2	390	5	2	390	5	2
22	392	5	2	392	5	2	391	5	2
23	394	5	2	394	5	2	393	5	2
24	395	5	2	395	5	2	394	5	2
25	397	5	2	397	5	2	396	5	2
26	399	5	2	398	5	2	397	5	2
27	400	5	3	400	5	3	399	5	2
28	402	5	3	402	5	3	400	5	3
29	403	5	3	403	5	3	402	5	3
30	405	5	3	405	5	3	404	5	3
31	406	5	3	406	5	3	405	5	3
32	408	5	3	408	5	3	407	5	3
33	410	5	3	410	5	3	409	5	3
34	411	5	3	411	5	3	410	5	3
35	413	5	3	413	5	3	412	5	3
36	415	5	3	415	5	3	414	5	3
37	417	5	3	417	5	3	416	5	3
38	419	6	3	419	5	3	418	6	3
39	421	6	4	421	6	4	420	6	4
40	424	6	4	423	6	4	423	6	4
41	426	6	4	426	6	4	425	6	4
42	429	6	4	428	6	4	428	7	4
43	432	7	4	431	7	4	431	7	4
44	435	7	4	435	7	4	435	7	4
45	439	8	4	438	8	4	439	8	4
46	444	9	4	443	9	4	444	9	4
47	450	10	4	449	10	4	450	10	4
48	457	12	4	456	12	4	458	12	4
49	469	16	4	468	16	4	470	16	4
50	489	25	4	488	25	4	490	25	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate the achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Table E.3. Algebra I Forms

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
0	329	24	1	332	23	1	331	23	1
1	345	13	1	348	13	1	347	13	1
2	355	10	1	357	9	1	357	9	1
3	360	8	1	363	8	1	362	8	1
4	365	7	1	367	7	1	367	7	1
5	368	6	1	371	6	1	370	6	1
6	372	6	1	374	6	1	373	6	1
7	374	6	1	376	6	1	375	6	1
8	377	5	1	379	5	1	378	5	1
9	379	5	1	381	5	1	380	5	1
10	381	5	1	382	5	1	382	5	1
11	382	5	1	384	5	1	383	5	1
12	384	5	1	386	4	1	385	5	1
13	386	4	1	387	4	1	387	4	1
14	387	4	1	389	4	2	388	4	1
15	389	4	2	390	4	2	390	4	2
16	390	4	2	392	4	2	391	4	2
17	391	4	2	393	4	2	392	4	2
18	393	4	2	394	4	2	394	4	2
19	394	4	2	395	4	2	395	4	2
20	395	4	2	397	4	2	396	4	2
21	396	4	2	398	4	2	397	4	2
22	398	4	2	399	4	2	398	4	2
23	399	4	2	400	4	3	400	4	3
24	400	4	3	401	4	3	401	4	3
25	401	4	3	402	4	3	402	4	3
26	402	4	3	403	4	3	403	4	3
27	404	4	3	404	4	3	404	4	3
28	405	4	3	405	4	3	406	4	3
29	406	4	3	407	4	3	407	4	3
30	407	4	3	408	4	3	408	4	3
31	408	4	3	409	4	4	409	4	4
32	410	4	4	410	4	4	410	4	4
33	411	4	4	411	4	4	412	4	4
34	412	4	4	412	4	4	413	4	4
35	414	4	4	414	4	4	414	4	4
36	415	4	4	415	4	4	416	4	4
37	417	4	4	416	4	4	417	4	4

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
38	418	5	4	418	4	4	419	4	4
39	420	5	4	419	5	4	420	5	4
40	422	5	4	421	5	4	422	5	4
41	423	5	4	423	5	4	424	5	4
42	425	5	4	425	5	4	426	5	4
43	428	5	4	427	5	4	428	5	4
44	430	6	4	429	6	4	430	6	4
45	433	6	4	432	6	4	433	6	4
46	436	7	4	436	7	4	436	7	4
47	441	8	4	440	8	4	440	8	4
48	446	9	4	446	9	4	446	9	4
49	456	13	4	455	13	4	455	13	4
50	471	23	4	471	24	4	471	23	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate the achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Table E.4. Algebra II Forms

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
0	335	21	1	334	22	1	335	22	1
1	350	12	1	349	12	1	350	12	1
2	358	9	1	357	9	1	359	9	1
3	363	7	1	363	7	1	365	7	1
4	367	6	1	367	6	1	369	6	1
5	370	6	1	370	6	1	372	6	1
6	373	5	1	372	5	1	375	5	1
7	375	5	1	375	5	1	377	5	1
8	377	5	1	377	5	1	379	5	1
9	379	5	1	379	5	1	381	5	1
10	381	4	1	380	4	1	383	4	1
11	383	4	1	382	4	1	384	4	1
12	384	4	1	383	4	1	386	4	1
13	386	4	1	385	4	1	387	4	1
14	387	4	1	386	4	1	388	4	2
15	389	4	2	388	4	2	390	4	2
16	390	4	2	389	4	2	391	4	2
17	391	4	2	390	4	2	392	4	2
18	393	4	2	391	4	2	393	4	2
19	394	4	2	393	4	2	394	4	2
20	395	4	2	394	4	2	395	4	2

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
21	396	4	2	395	4	2	396	4	2
22	397	4	2	396	4	2	398	4	2
23	399	4	2	397	4	2	399	4	2
24	400	4	3	398	4	2	400	4	3
25	401	4	3	399	3	2	401	4	3
26	402	4	3	400	3	3	402	4	3
27	403	4	3	401	3	3	403	4	3
28	405	4	3	402	3	3	404	4	3
29	406	4	3	403	3	3	405	4	3
30	407	4	3	404	3	3	406	4	3
31	408	4	3	405	3	3	407	4	3
32	409	4	3	406	3	3	408	4	3
33	411	4	4	407	4	3	410	4	3
34	412	4	4	408	4	3	411	4	4
35	413	4	4	410	4	3	412	4	4
36	414	4	4	411	4	4	413	4	4
37	416	4	4	412	4	4	415	4	4
38	417	4	4	413	4	4	416	4	4
39	419	4	4	414	4	4	418	4	4
40	420	4	4	416	4	4	419	4	4
41	422	5	4	417	4	4	421	5	4
42	424	5	4	419	5	4	423	5	4
43	426	5	4	421	5	4	425	5	4
44	429	5	4	423	5	4	428	5	4
45	431	6	4	426	6	4	430	6	4
46	435	6	4	429	6	4	434	6	4
47	439	7	4	433	7	4	438	7	4
48	444	9	4	438	9	4	443	9	4
49	453	12	4	447	12	4	452	12	4
50	467	22	4	461	22	4	467	22	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate the achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Table E.5. Geometry Forms

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
0	325	25	1	329	25	1	332	25	1
1	339	14	1	346	14	1	348	14	1
2	349	10	1	356	10	1	358	10	1
3	356	9	1	362	8	1	364	8	1

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
4	360	8	1	366	7	1	368	7	1
5	364	7	1	370	7	1	372	6	1
6	367	6	1	373	6	1	374	6	1
7	370	6	1	376	6	1	377	6	1
8	373	6	1	378	6	1	379	5	1
9	375	5	1	380	5	1	381	5	1
10	377	5	1	382	5	1	383	5	1
11	379	5	1	384	5	1	385	5	1
12	381	5	1	386	5	1	386	5	1
13	383	5	1	388	5	2	388	4	2
14	384	5	1	389	5	2	389	4	2
15	386	5	1	391	4	2	390	4	2
16	388	5	2	392	4	2	392	4	2
17	389	5	2	394	4	2	393	4	2
18	391	4	2	395	4	2	394	4	2
19	392	4	2	396	4	2	395	4	2
20	394	4	2	398	4	2	397	4	2
21	395	4	2	399	4	2	398	4	2
22	396	4	2	400	4	3	399	4	2
23	398	4	2	401	4	3	400	4	3
24	399	4	2	403	4	3	401	4	3
25	401	4	3	404	4	3	402	4	3
26	402	4	3	405	4	3	404	4	3
27	404	4	3	406	4	3	405	4	3
28	405	4	3	407	4	3	406	4	3
29	406	4	3	409	4	3	407	4	3
30	408	4	3	410	4	3	408	4	3
31	409	5	3	411	4	3	410	4	3
32	411	5	3	412	4	3	411	4	3
33	412	5	3	414	4	4	412	4	3
34	414	5	4	415	4	4	414	4	4
35	416	5	4	416	4	4	415	4	4
36	417	5	4	418	4	4	416	4	4
37	419	5	4	419	5	4	418	5	4
38	421	5	4	421	5	4	419	5	4
39	423	5	4	423	5	4	421	5	4
40	425	5	4	424	5	4	423	5	4
41	427	5	4	426	5	4	425	5	4
42	430	6	4	429	6	4	427	5	4
43	432	6	4	431	6	4	429	6	4

RS	Fall 2018 Core A			Spring 2019 Core C			Spring 2019 Core D		
	SS	CSEM	PL	SS	CSEM	PL	SS	CSEM	PL
44	435	6	4	434	6	4	432	6	4
45	438	7	4	437	7	4	435	7	4
46	442	7	4	441	8	4	438	7	4
47	446	8	4	446	9	4	443	8	4
48	452	10	4	452	10	4	449	10	4
49	462	14	4	463	14	4	459	14	4
50	479	25	4	480	25	4	475	25	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate the achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Table E.6. Biology Forms

RS	Fall 2018 & Spring 2019 Core A			Fall 2018 & Spring 2019 Core B		
	SS	CSEM	PL	SS	CSEM	PL
0	325	25	1	325	25	1
1	330	14	1	333	14	1
2	340	10	1	343	10	1
3	347	8	1	349	8	1
4	351	7	1	353	7	1
5	355	7	1	356	7	1
6	358	6	1	359	6	1
7	361	6	1	362	6	1
8	363	5	1	364	5	1
9	365	5	1	366	5	1
10	367	5	1	368	5	1
11	369	5	1	370	5	1
12	370	5	1	371	5	1
13	372	5	1	373	4	1
14	373	4	1	374	4	1
15	375	4	1	376	4	1
16	376	4	1	377	4	1
17	377	4	1	378	4	1
18	379	4	1	380	4	1
19	380	4	1	381	4	2
20	381	4	2	382	4	2
21	383	4	2	383	4	2
22	384	4	2	384	4	2
23	385	4	2	386	4	2
24	386	4	2	387	4	2
25	387	4	2	388	4	2

RS	Fall 2018 & Spring 2019 Core A			Fall 2018 & Spring 2019 Core B		
	SS	CSEM	PL	SS	CSEM	PL
26	389	4	2	389	4	2
27	390	4	2	390	4	2
28	391	4	2	391	4	2
29	392	4	2	393	4	2
30	394	4	2	394	4	2
31	395	4	2	395	4	2
32	396	4	2	396	4	2
33	398	4	2	398	4	2
34	399	4	2	399	4	2
35	400	4	3	400	4	3
36	402	4	3	402	4	3
37	403	5	3	403	5	3
38	405	5	3	405	5	3
39	406	5	3	406	5	3
40	408	5	3	408	5	3
41	410	5	3	410	5	3
42	412	5	4	412	5	4
43	414	6	4	414	6	4
44	417	6	4	417	6	4
45	419	6	4	420	7	4
46	423	7	4	423	7	4
47	427	8	4	428	8	4
48	433	10	4	434	10	4
49	443	14	4	444	14	4
50	459	25	4	461	25	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Table E.7. Physical Science Forms

RS	Fall 2018 & Spring 2019 Core A		
	SS	CSEM	PL
0	325	25	1
1	325	17	1
2	332	13	1
3	340	11	1
4	345	9	1
5	350	8	1
6	354	8	1
7	357	7	1

RS	Fall 2018 & Spring 2019 Core A		
	SS	CSEM	PL
8	360	7	1
9	363	7	1
10	365	6	1
11	368	6	1
12	370	6	1
13	372	6	1
14	374	6	1
15	376	6	1
16	378	5	1
17	379	5	1
18	381	5	1
19	383	5	2
20	384	5	2
21	386	5	2
22	387	5	2
23	389	5	2
24	390	5	2
25	392	5	2
26	393	5	2
27	395	5	2
28	396	5	2
29	397	5	2
30	399	5	2
31	400	5	3
32	402	5	3
33	404	5	3
34	405	5	3
35	407	5	3
36	409	6	3
37	410	6	3
38	412	6	3
39	414	6	3
40	417	6	4
41	419	7	4
42	422	7	4
43	425	7	4
44	428	8	4
45	432	9	4
46	437	10	4

RS	Fall 2018 & Spring 2019 Core A		
	SS	CSEM	PL
47	443	11	4
48	451	13	4
49	465	18	4
50	487	25	4

Note. The column names indicate the raw score (RS), the scaled score (SS), the conditional standard error of measurement (CSEM), and the achievement/performance level (PL). The values of 1 through 4 in the PL column of the tables indicate the achievement levels (Below-Basic, Basic, Proficient, and Advanced) respectively.

Appendix F: Descriptive Statistics by Demographic Group

Table F.1. Scale Score Descriptive Statistics by Demographic Group—Gender, Fall 2018

Test Period	Content Area	Gender	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	Female	76	363	435	402.50	14.48
		Male	88	330	429	395.76	16.48
	English II	Female	1,203	325	444	397.24	16.28
		Male	1,330	325	444	391.08	17.10
	Algebra I	Female	2,508	329	471	396.80	14.97
		Male	2,769	329	471	395.56	15.60
	Algebra II	Female	272	358	432	395.63	13.63
		Male	272	352	465	397.11	16.45
	Geometry	Female	72	377	462	410.10	16.61
		Male	66	381	446	414.47	14.90
	Biology	Female	1,233	330	461	391.17	16.85
		Male	1,270	325	459	388.68	17.29
	Physical Science	Female	22	365	399	383.64	9.73
		Male	17	365	407	386.18	12.10

Table F.2. Scale Score Descriptive Statistics by Demographic Group—Gender, Spring 2019

Test Period	Content Area	Gender	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	Female	5,524	325	456	404.56	13.95
		Male	5,567	325	456	399.38	15.15
	English II	Female	30,368	325	468	404.55	13.98
		Male	31,076	325	458	400.03	15.08
	Algebra I	Female	29,541	331	471	398.91	12.49
		Male	30,042	332	471	398.21	13.22
	Algebra II	Female	8,313	335	438	399.31	10.65
		Male	7,286	334	443	400.18	11.49
	Geometry	Female	1,847	366	459	398.93	11.78
		Male	1,715	366	459	399.74	12.96
	Biology	Female	29,899	325	461	395.65	14.36
		Male	30,478	325	461	395.27	15.61
	Physical Science	Female	1,105	354	465	394.55	14.05
		Male	1,181	350	451	396.06	15.44

Table F.3. Scale Score Descriptive Statistics by Demographic Group—Ethnicity, Fall 2018

Test Period	Content Area	Ethnicity	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	American Indian/Alaskan Native	1	--	--	--	--
		Asian	6	--	--	--	--
		Black (not Hispanic)	33	375	429	402.42	13.56
		Hispanic	28	330	429	389.61	18.84
		Multi-racial	4	--	--	--	--
		White (not Hispanic)	92	372	435	399.84	14.74
	English II	American Indian/Alaskan Native	9	--	--	--	--
		Asian	52	365	444	399.98	17.62
		Black (not Hispanic)	546	351	426	385.97	13.16
		Hispanic	278	325	429	392.27	15.66
		Multi-racial	91	351	432	390.14	16.36
		Pacific Islander	9	--	--	--	--
		White (not Hispanic)	1,548	325	444	397.18	17.42
	Algebra I	American Indian/Alaskan Native	22	329	425	396.73	20.97
		Asian	88	377	471	410.19	19.97
		Black (not Hispanic)	1,029	355	446	387.72	11.00
		Hispanic	449	329	430	391.72	13.10
		Multi-racial	178	345	441	393.80	14.84
		Pacific Islander	11	368	418	388.73	13.50
		White (not Hispanic)	3,498	329	471	398.99	15.33
	Algebra II	Asian	11	376	465	406.45	22.74
		Black (not Hispanic)	45	366	410	386.07	13.51
		Hispanic	50	366	417	390.76	13.26
		Multi-racial	21	352	436	395.38	18.36
		White (not Hispanic)	417	358	441	397.94	14.46
	Geometry	Asian	5	--	--	--	--
		Black (not Hispanic)	8	--	--	--	--
		Hispanic	10	377	421	396.20	14.44
		Multi-racial	4	--	--	--	--
		White (not Hispanic)	111	379	462	413.19	15.56
	Biology	American Indian/Alaskan Native	13	365	433	395.00	18.84
		Asian	76	359	459	401.01	18.52
		Black (not Hispanic)	472	333	427	378.95	12.06
		Hispanic	216	358	434	386.80	15.21
		Multi-racial	113	355	461	391.89	17.39
		Pacific Islander	6	--	--	--	--
		White (not Hispanic)	1,595	325	459	392.99	17.03

Test Period	Content Area	Ethnicity	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	Physical Science	Black (not Hispanic)	1	--	--	--	--
		Hispanic	12	365	396	379.75	8.64
		White (not Hispanic)	26	365	407	387.19	11.15

Table F.4. Scale Score Descriptive Statistics by Demographic Group—Ethnicity, Spring 2019

Test Period	Content Area	Ethnicity	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	American Indian/Alaskan Native	38	372	419	397.47	13.45
		Asian	133	353	445	404.55	16.41
		Black (not Hispanic)	1,428	325	435	392.86	17.22
		Hispanic	715	325	435	398.17	15.67
		Multi-racial	269	359	428	401.10	13.06
		Pacific Islander	32	364	435	400.72	12.89
		White (not Hispanic)	8,476	325	456	403.82	13.63
	English II	American Indian/Alaskan Native	224	367	443	401.58	13.27
		Asian	1,385	352	456	408.31	15.11
		Black (not Hispanic)	8,685	325	443	394.07	14.34
		Hispanic	3,881	325	444	399.03	14.33
		Multi-racial	1,938	325	449	402.63	14.48
		Pacific Islander	148	361	426	396.26	12.69
		White (not Hispanic)	45,181	325	468	403.94	14.22
	Algebra I	American Indian/Alaskan Native	236	371	430	395.23	12.47
		Asian	1,349	367	471	408.65	15.12
		Black (not Hispanic)	8,298	332	455	391.70	11.34
		Hispanic	3,898	348	446	396.31	12.00
		Multi-racial	2,057	332	455	399.08	12.78
		Pacific Islander	170	367	440	395.36	11.31
		White (not Hispanic)	43,568	331	471	399.76	12.61
	Algebra II	American Indian/Alaskan Native	38	382	423	399.71	10.45
		Asian	629	367	443	407.20	12.07
		Black (not Hispanic)	1,133	367	438	393.74	10.27
		Hispanic	826	370	434	397.88	10.55
		Multi-racial	501	369	429	398.87	11.06
		Pacific Islander	21	365	430	396.10	17.03
		White (not Hispanic)	12,450	334	438	400.04	10.80
	Geometry	American Indian/Alaskan Native	7	--	--	--	--
		Asian	72	386	459	418.81	15.20
		Black (not Hispanic)	122	373	418	396.67	10.20
		Hispanic	161	372	432	396.72	10.69

Test Period	Content Area	Ethnicity	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	Geometry	Multi-racial	87	376	434	398.98	10.59
		Pacific Islander	6	--	--	--	--
		White (not Hispanic)	3,107	366	459	399.12	12.12
	Biology	American Indian/ Alaskan Native	227	359	434	394.21	13.74
		Asian	1,326	356	461	403.28	16.57
		Pacific Islander	8,866	325	459	386.00	13.45
		Black (not Hispanic)	3,779	349	444	391.92	14.29
		Hispanic	1,954	325	444	395.55	14.71
		White (not Hispanic)	160	363	423	388.54	12.66
		Multi-racial	44,002	325	461	397.48	14.49
	Physical Science	American Indian/ Alaskan Native	8	--	--	--	--
		Asian	15	368	422	397.27	17.01
		Pacific Islander	120	357	419	385.62	13.02
		Black (not Hispanic)	54	365	437	391.94	13.57
		Hispanic	53	357	425	390.36	14.77
		White (not Hispanic)	2	--	--	--	--
		Multi-racial	2,032	350	465	396.13	14.70

Table F.5. Scale Score Descriptive Statistics by Demographic Group—Migrant Status, Fall 2018

Test Period	Content Area	Migrant	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	No	163	330	435	399.01	15.87
		Yes	1	--	--	--	--
	English II	No	2,540	325	444	394.00	17.00
		Yes	2	--	--	--	--
	Algebra I	No	5,310	329	471	396.11	15.30
		Yes	2	--	--	--	--
	Algebra II	No	544	352	465	396.34	15.16
		Yes	2	--	--	--	--
	Geometry	No	137	377	462	412.29	15.92
		Yes	1	--	--	--	--
	Biology	No	2,501	325	461	389.89	17.16
		Yes	4	--	--	--	--
	Physical Science	No	38	365	407	384.97	10.80
		Yes	1	--	--	--	--

Table F.6. Scale Score Descriptive Statistics by Demographic Group—Migrant Status, Spring 2019

Test Period	Content Area	Migrant	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	No	11,095	325	456	401.96	14.79
		Yes	7	--	--	--	--
	English II	No	61,543	325	468	402.26	14.72
		Yes	16	370	415	386.94	13.86
	Algebra I	No	59,743	331	471	398.54	12.87
		Yes	19	374	416	389.74	10.96
	Algebra II	No	15,635	334	443	399.71	11.05
		Yes	2	--	--	--	--
	Geometry	No	3,575	366	459	399.35	12.38
		Yes	1	--	--	--	--
	Biology	No	60,405	325	461	395.46	15.00
		Yes	20	367	406	382.20	10.35
	Physical Science	No	2,285	350	465	395.34	14.80
		Yes	1	--	--	--	--

Table F.7. Scale Score Descriptive Statistics by Demographic Group—Free and Reduced Lunch, Fall 2018

Test Period	Content Area	FRL	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	No	46	372	429	406.35	12.65
		Yes	118	330	435	395.97	16.13
	English II	No	1,175	325	444	399.09	18.07
		Yes	1,367	351	435	389.61	14.65
	Algebra I	No	2,822	329	471	401.63	15.56
		Yes	2,490	329	446	389.83	12.29
	Algebra II	No	376	358	465	398.65	14.74
		Yes	170	352	432	390.96	14.90
	Geometry	No	90	386	462	416.94	13.56
		Yes	48	377	435	403.27	16.27
	Biology	No	1,305	325	461	396.39	17.44
		Yes	1,200	330	459	382.81	13.72
	Physical Science	No	--	--	--	--	--
		Yes	39	365	407	384.74	10.75

Table F.8. Scale Score Descriptive Statistics by Demographic Group—Free and Reduced Lunch, Spring 2019

Test Period	Content Area	FRL	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	No	5,054	325	456	406.35	12.90
		Yes	6,048	325	449	398.27	15.26
	English II	No	35,008	325	468	406.10	13.60
		Yes	26,551	325	449	397.17	14.62
	Algebra I	No	33,023	331	471	402.17	12.72
		Yes	26,739	331	455	394.05	11.58
	Algebra II	No	11,452	334	443	401.21	10.93
		Yes	4,185	334	438	395.60	10.33
	Geometry	No	2,244	366	459	401.36	13.01
		Yes	1,332	366	437	395.97	10.37
	Biology	No	34,321	325	461	399.61	14.29
		Yes	26,104	325	459	390.00	14.14
	Physical Science	No	1,292	350	465	398.74	14.69
		Yes	994	350	451	390.90	13.73

Table F.9. Scale Score Descriptive Statistics by Demographic Group—Limited English Proficient, Fall 2018

Test Period	Content Area	LEP	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	No	132	368	435	400.47	14.89
		Yes	32	330	421	392.34	18.36
	English II	No	2,352	325	444	394.66	17.12
		Yes	190	325	417	385.76	12.75
	Algebra I	No	5,074	329	471	396.51	15.33
		Yes	238	329	428	387.46	11.76
	Algebra II	No	533	352	465	396.46	15.19
		Yes	13	372	408	388.08	13.50
	Geometry	No	138	377	462	412.19	15.91
		Yes	--	--	--	--	--
	Biology	No	2,393	325	461	390.33	17.27
		Yes	112	359	414	380.38	10.86
	Physical Science	No	35	365	407	385.97	10.45
		Yes	4	--	--	--	--

Table F.10. Scale Score Descriptive Statistics by Demographic Group—Limited English Proficient, Spring 2019

Test Period	Content Area	LEP	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	No	10,697	325	456	402.36	14.60
		Yes	405	325	435	391.06	15.65
	English II	No	59,720	325	468	402.62	14.62
		Yes	1,839	325	435	390.37	12.99
	Algebra I	No	57,573	331	471	398.73	12.87
		Yes	2,189	348	440	393.48	11.94
	Algebra II	No	15,382	334	443	399.78	11.02
		Yes	255	365	438	395.23	12.08
	Geometry	No	3,496	366	459	399.45	12.41
		Yes	80	373	419	395.21	9.98
	Biology	No	58,498	325	461	395.83	14.92
		Yes	1,927	349	434	384.34	13.05
	Physical Science	No	2,267	350	465	395.39	14.80
		Yes	19	365	419	388.26	13.57

Table F.11. Scale Score Descriptive Statistics by Demographic Group—Title I, Fall 2018

Test Period	Content Area	Title I	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	No	10	372	393	380.20	7.48
		Yes	154	330	435	400.10	15.55
	English II	No	2,020	325	444	394.62	17.42
		Yes	522	345	432	391.57	14.98
	Algebra I	No	4,570	329	471	397.30	15.56
		Yes	742	355	436	388.74	11.09
	Algebra II	No	479	352	465	397.91	14.89
		Yes	67	363	409	384.46	11.82
	Geometry	No	113	386	462	416.78	13.14
		Yes	25	377	417	391.44	9.52
	Biology	No	2,127	325	461	391.54	17.49
		Yes	378	355	419	380.57	11.32
	Physical Science	No	--	--	--	--	--
		Yes	39	365	407	384.74	10.75

Table F.12. Scale Score Descriptive Statistics by Demographic Group—Title I, Spring 2019

Test Period	Content Area	Title I	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	No	9,149	325	456	403.46	13.71
		Yes	1,953	325	439	394.87	17.38
	English II	No	56,222	325	468	403.01	14.52
		Yes	5,337	325	456	394.24	14.52
	Algebra I	No	53,561	331	471	399.24	12.82
		Yes	6,201	348	446	392.47	11.71
	Algebra II	No	14,930	334	443	399.99	11.02
		Yes	707	367	426	393.69	9.97
	Geometry	No	3,374	366	459	399.51	12.45
		Yes	202	366	434	396.71	10.77
	Biology	No	54,586	325	461	396.39	14.83
		Yes	5,839	325	444	386.80	13.82
	Physical Science	No	2,168	350	465	395.51	14.91
		Yes	118	365	428	391.95	12.21

Table F.13. Scale Score Descriptive Statistics by Demographic Group—Students with IEPs, Fall 2018

Test Period	Content Area	IEP	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	No	156	330	435	399.58	15.87
		Yes	8	--	--	--	--
	English II	No	2,241	325	444	395.80	16.63
		Yes	301	325	417	380.53	13.16
	Algebra I	No	4,707	329	471	397.46	15.14
		Yes	605	329	433	385.49	12.05
	Algebra II	No	537	352	465	396.33	15.20
		Yes	9	--	--	--	--
	Geometry	No	136	377	462	412.17	15.81
		Yes	2	--	--	--	--
	Biology	No	2,201	325	461	391.37	17.03
		Yes	304	325	433	379.13	14.01
	Physical Science	No	38	365	407	385.13	10.62
		Yes	1	--	--	--	--

Table F.14. Scale Score Descriptive Statistics by Demographic Group—Students with IEPs, Spring 2019

Test Period	Content Area	IEP	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	No	10,088	325	456	403.42	14.08
		Yes	1,014	325	431	387.35	13.75
	English II	No	54,930	325	468	404.00	13.89
		Yes	6,629	325	443	387.73	13.42
	Algebra I	No	53,242	331	471	399.89	12.45
		Yes	6,520	331	455	387.49	10.76
	Algebra II	No	15,415	334	443	399.80	11.01
		Yes	222	369	430	393.26	12.30
	Geometry	No	3,426	366	459	399.71	12.33
		Yes	150	366	421	391.27	10.59
	Biology	No	54,004	325	461	396.96	14.48
		Yes	6,421	325	444	382.82	13.29
	Physical Science	No	2,113	354	465	396.50	14.46
		Yes	173	350	422	381.01	10.88

Table F.15. Scale Score Descriptive Statistics by Demographic Group—Students with Accommodations, Fall 2018

Test Period	Content Area	Accom.	*n*-Count	Min.	Max.	Mean	SD
Fall 2018	English I	No	145	330	435	400.20	16.05
		Yes	19	372	406	388.84	10.38
	English II	No	2,274	325	444	395.24	16.91
		Yes	268	325	435	383.41	13.67
	Algebra I	No	4,846	329	471	396.91	15.33
		Yes	466	329	433	387.70	12.18
	Algebra II	No	540	352	465	396.26	15.19
		Yes	6	--	--	--	--
	Geometry	No	137	377	462	412.12	15.95
		Yes	1	--	--	--	--
	Biology	No	2,251	325	461	390.99	17.27
		Yes	254	351	423	380.11	12.50
	Physical Science	No	39	365	407	384.74	10.75
		Yes	--	--	--	--	--

Table F.16. Scale Score Descriptive Statistics by Demographic Group—Students with Accommodations, Spring 2019

Test Period	Content Area	Accom.	*n*-Count	Min.	Max.	Mean	SD
Spring 2019	English I	No	9,798	325	456	402.95	14.44
		Yes	1,304	325	449	394.43	15.24
	English II	No	52,838	325	458	403.67	14.13
		Yes	8,721	325	468	393.66	15.34
	Algebra I	No	51,695	331	471	399.78	12.57
		Yes	8,067	332	455	390.59	11.92
	Algebra II	No	14,738	334	443	399.87	11.06
		Yes	899	369	433	397.13	10.60
	Geometry	No	3,416	366	459	399.68	12.35
		Yes	160	366	421	392.36	10.82
	Biology	No	51,590	325	461	396.82	14.62
		Yes	8,835	325	459	387.54	14.72
	Physical Science	No	2,123	354	465	396.33	14.51
		Yes	163	350	422	382.31	12.14

Appendix G: Achievement-Level Distributions by Demographic Group

Table G.1. Achievement-Level Distributions—Gender, Fall 2018

Test Period	Content Area	Gender	Achievement Level	Freq.	%
Fall 2018	English I	Female	Below Basic	7	9.21
			Basic	20	26.32
			Proficient	35	46.05
			Advanced	14	18.42
			Proficient + Advanced	49	64.47
			Total	76	100.00
		Male	Below Basic	21	23.86
			Basic	30	34.09
			Proficient	26	29.55
			Advanced	11	12.50
			Proficient + Advanced	37	42.05
			Total	88	100.00
	English II	Female	Below Basic	272	22.63
			Basic	359	29.87
			Proficient	486	40.43
			Advanced	85	7.07
			Proficient + Advanced	571	47.50
			Total	1,202	100.00
		Male	Below Basic	523	39.32
			Basic	356	26.77
			Proficient	391	29.40
			Advanced	60	4.51
			Proficient + Advanced	451	33.91
			Total	1,330	100.00
	Algebra I	Female	Below Basic	777	30.98
			Basic	727	28.99
			Proficient	450	17.94
			Advanced	554	22.09
			Proficient + Advanced	1,004	40.03
			Total	2,508	100.00
		Male	Below Basic	1,011	36.52
			Basic	735	26.55
			Proficient	440	15.90
			Advanced	582	21.03
			Proficient + Advanced	1,022	36.92
			Total	2,768	100.00
	Algebra II	Female	Below Basic	80	29.41
			Basic	66	24.26
			Proficient	95	34.93
			Advanced	31	11.40
			Proficient + Advanced	126	46.32
			Total	272	100.00
		Male	Below Basic	82	30.15
			Basic	60	22.06
			Proficient	82	30.15
			Advanced	48	17.65
			Proficient + Advanced	130	47.79
			Total	272	100.00

Test Period	Content Area	Gender	Achievement Level	Freq.	%
Fall 2018	Geometry	Female	Below Basic	6	8.33
			Basic	14	19.44
			Proficient	22	30.56
			Advanced	30	41.67
			Proficient + Advanced	52	72.22
			Total	72	100.00
		Male	Below Basic	4	6.06
			Basic	8	12.12
			Proficient	18	27.27
			Advanced	36	54.55
			Proficient + Advanced	54	81.82
			Total	66	100.00
	Biology	Female	Below Basic	371	30.09
			Basic	476	38.61
			Proficient	229	18.57
			Advanced	157	12.73
			Proficient + Advanced	386	31.31
			Total	1,233	100.00
		Male	Below Basic	468	36.85
			Basic	457	35.98
			Proficient	201	15.83
			Advanced	144	11.34
			Proficient + Advanced	345	27.17
			Total	1,270	100.00
	Physical Science	Female	Below Basic	12	54.55
			Basic	10	45.45
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	22	100.00
		Male	Below Basic	6	35.29
			Basic	8	47.06
			Proficient	3	17.65
			Advanced	--	--
			Proficient + Advanced	3	17.65
			Total	17	100.00

Table G.2. Achievement-Level Distributions—Gender, Spring 2019

Test Period	Content Area	Gender	Achievement Level	Freq.	%
Spring 2019	English I	Female	Below Basic	378	6.84
			Basic	1,439	26.05
			Proficient	2,386	43.20
			Advanced	1,320	23.90
			Proficient + Advanced	3,706	67.10
			Total	5,523	100.00
		Male	Below Basic	833	14.97
			Basic	1,776	31.91
			Proficient	2,134	38.34
			Advanced	823	14.79
			Proficient + Advanced	2,957	53.13
			Total	5,566	100.00

Test Period	Content Area	Gender	Achievement Level	Freq.	%
Spring 2019	English II	Female	Below Basic	2,311	7.61
			Basic	8,051	26.52
			Proficient	15,903	52.39
			Advanced	4,088	13.47
			Proficient + Advanced	19,991	65.86
			Total	30,353	100.00
		Male	Below Basic	4,718	15.20
			Basic	9,635	31.03
			Proficient	13,990	45.06
			Advanced	2,706	8.72
			Proficient + Advanced	16,696	53.77
			Total	31,049	100.00
	Algebra I	Female	Below Basic	6,320	21.40
			Basic	9,344	31.64
			Proficient	7,305	24.74
			Advanced	6,562	22.22
			Proficient + Advanced	13,867	46.96
			Total	29,531	100.00
		Male	Below Basic	7,567	25.20
			Basic	8,946	29.79
			Proficient	6,842	22.79
			Advanced	6,673	22.22
			Proficient + Advanced	13,515	45.01
			Total	30,028	100.00
	Algebra II	Female	Below Basic	1,107	13.33
			Basic	3,205	38.59
			Proficient	2,661	32.04
			Advanced	1,333	16.05
			Proficient + Advanced	3,994	48.09
			Total	8,306	100.00
		Male	Below Basic	1,017	13.97
			Basic	2,469	33.91
			Proficient	2,356	32.35
			Advanced	1,440	19.77
			Proficient + Advanced	3,796	52.13
			Total	7,282	100.00
	Geometry	Female	Below Basic	256	13.86
			Basic	779	42.18
			Proficient	589	31.89
			Advanced	223	12.07
			Proficient + Advanced	812	43.96
			Total	1,847	100.00
		Male	Below Basic	246	14.34
			Basic	690	40.23
			Proficient	516	30.09
			Advanced	263	15.34
			Proficient + Advanced	779	45.42
			Total	1,715	100.00
	Biology	Female	Below Basic	4,512	15.09
			Basic	13,656	45.68
			Proficient	7,476	25.01
			Advanced	4,254	14.23
			Proficient + Advanced	11,730	39.23

Test Period	Content Area	Gender	Achievement Level	Freq.	%
Spring 2019	Biology	Female	Total	29,898	100.00
		Male	Below Basic	5,625	18.46
			Basic	12,760	41.87
			Proficient	7,300	23.95
			Advanced	4,789	15.72
			Proficient + Advanced	12,089	39.67
			Total	30,474	100.00
	Physical Science	Female	Below Basic	204	18.46
			Basic	522	47.24
			Proficient	295	26.70
			Advanced	84	7.60
			Proficient + Advanced	379	34.30
			Total	1,105	100.00
		Male	Below Basic	235	19.90
			Basic	475	40.22
			Proficient	343	29.04
			Advanced	128	10.84
			Proficient + Advanced	471	39.88
			Total	1,181	100.00

Table G.3. Achievement-Level Distributions—Ethnicity, Fall 2018

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	English I	American Indian/ Alaskan Native	Below Basic	1	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00
		Asian	Below Basic	--	--
			Basic	1	16.67
			Proficient	2	33.33
			Advanced	3	50.00
			Proficient + Advanced	5	83.33
			Total	6	100.00
		Pacific Islander	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Black (not Hispanic)	Below Basic	3	9.09
			Basic	10	30.30
			Proficient	14	42.42
			Advanced	6	18.18
			Proficient + Advanced	20	60.61
			Total	33	100.00
		Hispanic	Below Basic	8	28.57
			Basic	13	46.43
			Proficient	6	21.43
			Advanced	1	3.57
			Proficient + Advanced	7	25.00
			Total	28	100.00

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	English I	White (not Hispanic)	Below Basic	15	16.30
			Basic	26	28.26
			Proficient	36	39.13
			Advanced	15	16.30
			Proficient + Advanced	51	55.43
			Total	92	100.00
		Multi-racial	Below Basic	1	25.00
			Basic	--	--
			Proficient	3	75.00
			Advanced	--	--
			Proficient + Advanced	3	75.00
			Total	4	100.00
	English II	American Indian/ Alaskan Native	Below Basic	1	11.11
			Basic	4	44.44
			Proficient	3	33.33
			Advanced	1	11.11
			Proficient + Advanced	4	44.44
			Total	9	100.00
		Asian	Below Basic	7	13.46
			Basic	21	40.38
			Proficient	17	32.69
			Advanced	7	13.46
			Proficient + Advanced	24	46.15
			Total	52	100.00
		Pacific Islander	Below Basic	5	55.56
			Basic	3	33.33
			Proficient	1	11.11
			Advanced	--	--
			Proficient + Advanced	1	11.11
			Total	9	100.00
		Black (not Hispanic)	Below Basic	263	48.17
			Basic	178	32.60
			Proficient	103	18.86
			Advanced	2	0.37
			Proficient + Advanced	105	19.23
			Total	546	100.00
		Hispanic	Below Basic	86	30.94
			Basic	92	33.09
			Proficient	93	33.45
			Advanced	7	2.52
			Proficient + Advanced	100	35.97
			Total	278	100.00
		White (not Hispanic)	Below Basic	398	25.73
			Basic	389	25.15
			Proficient	634	40.98
			Advanced	126	8.14
			Proficient + Advanced	760	49.13
			Total	1,547	100.00
		Multi-racial	Below Basic	35	38.46
			Basic	28	30.77
			Proficient	26	28.57
			Advanced	2	2.20

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	English II	Multi-racial	Proficient + Advanced	28	30.77
			Total	91	100.00
	Algebra I	American Indian/ Alaskan Native	Below Basic	7	31.82
			Basic	5	22.73
			Proficient	3	13.64
			Advanced	7	31.82
			Proficient + Advanced	10	45.45
			Total	22	100.00
		Asian	Below Basic	17	19.32
			Basic	13	14.77
			Proficient	11	12.50
			Advanced	47	53.41
			Proficient + Advanced	58	65.91
			Total	88	100.00
		Pacific Islander	Below Basic	5	45.45
			Basic	5	45.45
			Proficient	--	--
			Advanced	1	9.09
			Proficient + Advanced	1	9.09
			Total	11	100.00
		Black (not Hispanic)	Below Basic	579	56.32
			Basic	311	30.25
			Proficient	91	8.85
			Advanced	47	4.57
			Proficient + Advanced	138	13.42
			Total	1,028	100.00
		Hispanic	Below Basic	195	43.43
			Basic	138	30.73
			Proficient	67	14.92
			Advanced	49	10.91
			Proficient + Advanced	116	25.84
			Total	449	100.00
		White (not Hispanic)	Below Basic	915	26.16
			Basic	938	26.82
			Proficient	685	19.58
			Advanced	960	27.44
			Proficient + Advanced	1,645	47.03
			Total	3,498	100.00
		Multi-racial	Below Basic	69	38.76
			Basic	51	28.65
			Proficient	33	18.54
			Advanced	25	14.04
			Proficient + Advanced	58	32.58
			Total	178	100.00
	Algebra II	American Indian/ Alaskan Native	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Asian	Below Basic	2	18.18
			Basic	1	9.09

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	Algebra II	Asian	Proficient	6	54.55
			Advanced	2	18.18
			Proficient + Advanced	8	72.73
			Total	11	100.00
		Pacific Islander	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Black (not Hispanic)	Below Basic	24	53.33
			Basic	10	22.22
			Proficient	11	24.44
			Advanced	--	--
			Proficient + Advanced	11	24.44
			Total	45	100.00
		Hispanic	Below Basic	21	42.00
			Basic	14	28.00
			Proficient	14	28.00
			Advanced	1	2.00
			Proficient + Advanced	15	30.00
			Total	50	100.00
		White (not Hispanic)	Below Basic	109	26.14
			Basic	95	22.78
			Proficient	139	33.33
			Advanced	74	17.75
			Proficient + Advanced	213	51.08
			Total	417	100.00
		Multi-racial	Below Basic	6	28.57
			Basic	6	28.57
			Proficient	7	33.33
			Advanced	2	9.52
			Proficient + Advanced	9	42.86
			Total	21	100.00
	Geometry	American Indian/ Alaskan Native	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Asian	Below Basic	--	--
			Basic	1	20.00
			Proficient	2	40.00
			Advanced	2	40.00
			Proficient + Advanced	4	80.00
			Total	5	100.00
		Pacific Islander	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	Geometry	Black (not Hispanic)	Below Basic	--	--
			Basic	2	25.00
			Proficient	2	25.00
			Advanced	4	50.00
			Proficient + Advanced	6	75.00
			Total	8	100.00
		Hispanic	Below Basic	3	30.00
			Basic	4	40.00
			Proficient	1	10.00
			Advanced	2	20.00
			Proficient + Advanced	3	30.00
			Total	10	100.00
		White (not Hispanic)	Below Basic	7	6.31
			Basic	15	13.51
			Proficient	34	30.63
			Advanced	55	49.55
			Proficient + Advanced	89	80.18
			Total	111	100.00
		Multi-racial	Below Basic	--	--
			Basic	--	--
			Proficient	1	25.00
			Advanced	3	75.00
			Proficient + Advanced	4	100.00
			Total	4	100.00
	Biology	American Indian/ Alaskan Native	Below Basic	3	23.08
			Basic	5	38.46
			Proficient	3	23.08
			Advanced	2	15.38
			Proficient + Advanced	5	38.46
			Total	13	100.00
		Asian	Below Basic	10	13.16
			Basic	30	39.47
			Proficient	10	13.16
			Advanced	26	34.21
			Proficient + Advanced	36	47.37
			Total	76	100.00
		Pacific Islander	Below Basic	5	83.33
			Basic	1	16.67
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	6	100.00
		Black (not Hispanic)	Below Basic	285	60.38
			Basic	159	33.69
			Proficient	22	4.66
			Advanced	6	1.27
			Proficient + Advanced	28	5.93
			Total	472	100.00
		Hispanic	Below Basic	87	40.28
			Basic	87	40.28
			Proficient	22	10.19
			Advanced	20	9.26

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	Biology	Hispanic	Proficient + Advanced	42	19.44
			Total	216	100.00
		White (not Hispanic)	Below Basic	410	25.71
			Basic	605	37.93
			Proficient	348	21.82
			Advanced	232	14.55
			Proficient + Advanced	580	36.36
			Total	1,595	100.00
		Multi-racial	Below Basic	31	27.43
			Basic	42	37.17
			Proficient	25	22.12
			Advanced	15	13.27
			Proficient + Advanced	40	35.40
			Total	113	100.00
	Physical Science	American Indian/ Alaskan Native	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Asian	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Pacific Islander	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Black (not Hispanic)	Below Basic	1	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00
		Hispanic	Below Basic	8	66.67
			Basic	4	33.33
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	12	100.00
		White (not Hispanic)	Below Basic	9	34.62
			Basic	14	53.85
			Proficient	3	11.54
			Advanced	--	--
			Proficient + Advanced	3	11.54
			Total	26	100.00

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Fall 2018	Physical Science	Multi-racial	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--

Table G.4. Achievement-Level Distributions—Ethnicity, Spring 2019

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Spring 2019	English I	American Indian/ Alaskan Native	Below Basic	7	18.42
			Basic	12	31.58
			Proficient	14	36.84
			Advanced	5	13.16
			Proficient + Advanced	19	50.00
			Total	38	100.00
		Asian	Below Basic	19	14.29
			Basic	26	19.55
			Proficient	50	37.59
			Advanced	38	28.57
			Proficient + Advanced	88	66.17
			Total	133	100.00
		Pacific Islander	Below Basic	2	6.25
			Basic	14	43.75
			Proficient	13	40.63
			Advanced	3	9.38
			Proficient + Advanced	16	50.00
			Total	32	100.00
		Black (not Hispanic)	Below Basic	365	25.58
			Basic	533	37.35
			Proficient	416	29.15
			Advanced	113	7.92
			Proficient + Advanced	529	37.07
			Total	1,427	100.00
		Hispanic	Below Basic	118	16.53
			Basic	229	32.07
			Proficient	272	38.10
			Advanced	95	13.31
			Proficient + Advanced	367	51.40
			Total	714	100.00
		White (not Hispanic)	Below Basic	673	7.94
			Basic	2,312	27.28
			Proficient	3,642	42.97
			Advanced	1,849	21.81
			Proficient + Advanced	5,491	64.78
			Total	8,476	100.00
		Multi-racial	Below Basic	27	10.04
			Basic	89	33.09
			Proficient	113	42.01
			Advanced	40	14.87
			Proficient + Advanced	153	56.88
			Total	269	100.00

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Spring 2019	English II	American Indian/ Alaskan Native	Below Basic	21	9.38
			Basic	78	34.82
			Proficient	109	48.66
			Advanced	16	7.14
			Proficient + Advanced	125	55.80
			Total	224	100.00
		Asia	Below Basic	102	7.36
			Basic	257	18.56
			Proficient	687	49.60
			Advanced	339	24.48
			Proficient + Advanced	1,026	74.08
			Total	1,385	100.00
		Pacific Islander	Below Basic	26	17.57
			Basic	59	39.86
			Proficient	58	39.19
			Advanced	5	3.38
			Proficient + Advanced	63	42.57
			Total	148	100.00
		Black (not Hispanic)	Below Basic	2,091	24.11
			Basic	3,448	39.75
			Proficient	2,862	33.00
			Advanced	273	3.15
			Proficient + Advanced	3,135	36.14
			Total	8,674	100.00
		Hispanic	Below Basic	575	14.83
			Basic	1,330	34.30
			Proficient	1,704	43.94
			Advanced	269	6.94
			Proficient + Advanced	1,973	50.88
			Total	3,878	100.00
		White (not Hispanic)	Below Basic	4,008	8.88
			Basic	11,954	26.47
			Proficient	23,516	52.08
			Advanced	5,677	12.57
			Proficient + Advanced	29,193	64.65
			Total	45,155	100.00
		Multi-racial	Below Basic	205	10.59
			Basic	559	28.87
			Proficient	957	49.43
			Advanced	215	11.11
			Proficient + Advanced	1,172	60.54
			Total	1,936	100.00
	Algebra I	American Indian/ Alaskan Native	Below Basic	74	31.36
			Basic	75	31.78
			Proficient	48	20.34
			Advanced	39	16.53
			Proficient + Advanced	87	36.86
			Total	236	100.00
		Asian	Below Basic	132	9.79
			Basic	226	16.75
			Proficient	313	23.20
			Advanced	678	50.26
			Proficient + Advanced	991	73.46

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Spring 2019	Algebra I	Asian	Total	1,349	100.00
		Pacific Islander	Below Basic	51	30.00
			Basic	59	34.71
			Proficient	41	24.12
			Advanced	19	11.18
			Proficient + Advanced	60	35.29
			Total	170	100.00
		Black (not Hispanic)	Below Basic	3,593	43.35
			Basic	2,673	32.25
			Proficient	1,330	16.05
			Advanced	692	8.35
			Proficient + Advanced	2,022	24.40
			Total	8,288	100.00
		Hispanic	Below Basic	1,091	27.99
			Basic	1,324	33.97
			Proficient	850	21.81
			Advanced	633	16.24
			Proficient + Advanced	1,483	38.05
			Total	3,898	100.00
		White (not Hispanic)	Below Basic	8,496	19.51
			Basic	13,281	30.49
			Proficient	11,073	25.42
			Advanced	10,705	24.58
			Proficient + Advanced	21,778	50.00
			Total	43,555	100.00
		Multi-racial	Below Basic	447	21.74
			Basic	650	31.61
			Proficient	491	23.88
			Advanced	468	22.76
			Proficient + Advanced	959	46.64
			Total	2,056	100.00
	Algebra II	American Indian/ Alaskan Native	Below Basic	4	10.53
			Basic	17	44.74
			Proficient	8	21.05
			Advanced	9	23.68
			Proficient + Advanced	17	44.74
			Total	38	100.00
		Asian	Below Basic	38	6.05
			Basic	111	17.68
			Proficient	212	33.76
			Advanced	267	42.52
			Proficient + Advanced	479	76.27
			Total	628	100.00
		Pacific Islander	Below Basic	8	38.10
			Basic	4	19.05
			Proficient	5	23.81
			Advanced	4	19.05
			Proficient + Advanced	9	42.86
			Total	21	100.00
		Black (not Hispanic)	Below Basic	326	28.82
			Basic	486	42.97
			Proficient	240	21.22

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Spring 2019	Algebra II	Black (not Hispanic)	Advanced	79	6.98
			Proficient + Advanced	319	28.21
			Total	1,131	100.00
		Hispanic	Below Basic	134	16.26
			Basic	334	40.53
			Proficient	249	30.22
			Advanced	107	12.99
			Proficient + Advanced	356	43.20
			Total	824	100.00
		White (not Hispanic)	Below Basic	1,534	12.33
			Basic	4,537	36.46
			Proficient	4,147	33.33
			Advanced	2,226	17.89
			Proficient + Advanced	6,373	51.21
			Total	12,444	100.00
		Multi-racial	Below Basic	80	15.97
			Basic	185	36.93
			Proficient	156	31.14
			Advanced	80	15.97
			Proficient + Advanced	236	47.11
			Total	501	100.00
	Geometry	American Indian/ Alaskan Native	Below Basic	4	57.14
			Basic	2	28.57
			Proficient	1	14.29
			Advanced	--	--
			Proficient + Advanced	1	14.29
			Total	7	100.00
		Asian	Below Basic	1	1.39
			Basic	5	6.94
			Proficient	20	27.78
			Advanced	46	63.89
			Proficient + Advanced	66	91.67
			Total	72	100.00
		Pacific Islander	Below Basic	--	--
			Basic	2	33.33
			Proficient	3	50.00
			Advanced	1	16.67
			Proficient + Advanced	4	66.67
			Total	6	100.00
		Black (not Hispanic)	Below Basic	17	13.93
			Basic	61	50.00
			Proficient	33	27.05
			Advanced	11	9.02
			Proficient + Advanced	44	36.07
			Total	122	100.00
		Hispanic	Below Basic	25	15.53
			Basic	84	52.17
			Proficient	39	24.22
			Advanced	13	8.07
			Proficient + Advanced	52	32.30
			Total	161	100.00

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Spring 2019	Geometry	White (not Hispanic)	Below Basic	447	14.39
			Basic	1,277	41.10
			Proficient	976	31.41
			Advanced	407	13.10
			Proficient + Advanced	1,383	44.51
			Total	3,107	100.00
		Multi-racial	Below Basic	8	9.20
			Basic	38	43.68
			Proficient	33	37.93
			Advanced	8	9.20
			Proficient + Advanced	41	47.13
			Total	87	100.00
	Biology	American Indian/ Alaskan Native	Below Basic	40	17.62
			Basic	106	46.70
			Proficient	55	24.23
			Advanced	26	11.45
			Proficient + Advanced	81	35.68
			Total	227	100.00
		Asian	Below Basic	127	9.58
			Basic	401	30.24
			Proficient	362	27.30
			Advanced	436	32.88
			Proficient + Advanced	798	60.18
			Total	1,326	100.00
		Pacific Islander	Below Basic	47	29.38
			Basic	79	49.38
			Proficient	27	16.88
			Advanced	7	4.38
			Proficient + Advanced	34	21.25
			Total	160	100.00
		Black (not Hispanic)	Below Basic	3,351	37.81
			Basic	4,041	45.59
			Proficient	1,117	12.60
			Advanced	354	3.99
			Proficient + Advanced	1,471	16.60
			Total	8,863	100.00
		Hispanic	Below Basic	867	22.94
			Basic	1,813	47.98
			Proficient	736	19.48
			Advanced	363	9.61
			Proficient + Advanced	1,099	29.08
			Total	3,779	100.00
		White (not Hispanic)	Below Basic	5,383	12.23
			Basic	19,067	43.33
			Proficient	11,991	27.25
			Advanced	7,560	17.18
			Proficient + Advanced	19,551	44.43
			Total	44,001	100.00
		Multi-racial	Below Basic	296	15.16
			Basic	879	45.01
			Proficient	481	24.63
			Advanced	297	15.21

Test Period	Content Area	Ethnicity	Achievement Level	Freq.	%
Spring 2019	Biology	Multi-racial	Proficient + Advanced	778	39.84
			Total	1,953	100.00
	Physical Science	American Indian/ Alaskan Native	Below Basic	2	25.00
			Basic	3	37.50
			Proficient	3	37.50
			Advanced	--	--
			Proficient + Advanced	3	37.50
			Total	8	100.00
		Asian	Below Basic	4	26.67
			Basic	5	33.33
			Proficient	2	13.33
			Advanced	4	26.67
			Proficient + Advanced	6	40.00
			Total	15	100.00
		Pacific Islander	Below Basic	--	--
			Basic	2	100.00
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	2	100.00
		Black (not Hispanic)	Below Basic	48	40.00
			Basic	52	43.33
			Proficient	19	15.83
			Advanced	1	0.83
			Proficient + Advanced	20	16.67
			Total	120	100.00
		Hispanic	Below Basic	12	22.22
			Basic	27	50.00
			Proficient	13	24.07
			Advanced	2	3.70
			Proficient + Advanced	15	27.78
			Total	54	100.00
		White (not Hispanic)	Below Basic	355	17.47
			Basic	886	43.60
			Proficient	588	28.94
			Advanced	203	9.99
			Proficient + Advanced	791	38.93
			Total	2,032	100.00
		Multi-racial	Below Basic	18	33.96
			Basic	20	37.74
			Proficient	13	24.53
			Advanced	2	3.77
			Proficient + Advanced	15	28.30
			Total	53	100.00

Table G.5. Achievement-Level Distributions—Migrant, Fall 2018

Test Period	Content Area	Migrant	Achievement Level	Freq.	%
Fall 2018	English I	No	Below Basic	27	16.56
			Basic	50	30.67
			Proficient	61	37.42
			Advanced	25	15.34
			Proficient + Advanced	86	52.76
			Total	163	100.00
		Yes	Below Basic	1	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00
	English II	No	Below Basic	799	31.47
			Basic	716	28.20
			Proficient	878	34.58
			Advanced	146	5.75
			Proficient + Advanced	1,024	40.33
			Total	2,539	100.00
		Yes	Below Basic	--	--
			Basic	2	100.00
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	2	100.00
	Algebra I	No	Below Basic	1,803	33.96
			Basic	1,474	27.76
			Proficient	894	16.84
			Advanced	1,138	21.44
			Proficient + Advanced	2,032	38.27
			Total	5,309	100.00
		Yes	Below Basic	2	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	2	100.00
	Algebra II	No	Below Basic	162	29.78
			Basic	126	23.16
			Proficient	177	32.54
			Advanced	79	14.52
			Proficient + Advanced	256	47.06
			Total	544	100.00
		Yes	Below Basic	2	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	2	100.00
	Geometry	No	Below Basic	10	7.30
			Basic	21	15.33
			Proficient	40	29.20

Test Period	Content Area	Migrant	Achievement Level	Freq.	%
Fall 2018	Geometry	No	Advanced	66	48.18
			Proficient + Advanced	106	77.37
			Total	137	100.00
		Yes	Below Basic	--	--
			Basic	1	100.00
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00
	Biology	No	Below Basic	838	33.51
			Basic	932	37.27
			Proficient	430	17.19
			Advanced	301	12.04
			Proficient + Advanced	731	29.23
			Total	2,501	100.00
		Yes	Below Basic	2	50.00
			Basic	1	25.00
			Proficient	1	25.00
			Advanced	--	--
			Proficient + Advanced	1	25.00
			Total	4	100.00
	Physical Science	No	Below Basic	17	44.74
			Basic	18	47.37
			Proficient	3	7.89
			Advanced	--	--
			Proficient + Advanced	3	7.89
			Total	38	100.00
		Yes	Below Basic	1	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00

Table G.6. Achievement-Level Distributions—Migrant, Spring 2019

Test Period	Content Area	Migrant	Achievement Level	Freq.	%
Spring 2019	English I	No	Below Basic	1,211	10.92
			Basic	3,218	29.01
			Proficient	4,520	40.75
			Advanced	2,144	19.33
			Proficient + Advanced	6,664	60.07
			Total	11,093	100.00
		Yes	Below Basic	2	28.57
			Basic	3	42.86
			Proficient	1	14.29
			Advanced	1	14.29
			Proficient + Advanced	2	28.57
			Total	7	100.00
	English II	No	Below Basic	7,056	11.47
			Basic	17,713	28.80
			Proficient	29,925	48.66
			Advanced	6,804	11.06

Test Period	Content Area	Migrant	Achievement Level	Freq.	%
Spring 2019	English II	No	Proficient + Advanced	36,729	59.72
			Total	61,498	100.00
		Yes	Below Basic	7	43.75
			Basic	6	37.50
			Proficient	3	18.75
			Advanced	--	--
			Proficient + Advanced	3	18.75
			Total	16	100.00
	Algebra I	No	Below Basic	13,959	23.38
			Basic	18,339	30.71
			Proficient	14,171	23.73
			Advanced	13,248	22.18
			Proficient + Advanced	27,419	45.91
			Total	59,717	100.00
		Yes	Below Basic	10	52.63
			Basic	5	26.32
			Proficient	3	15.79
			Advanced	1	5.26
			Proficient + Advanced	4	21.05
			Total	19	100.00
	Algebra II	No	Below Basic	2,127	13.62
			Basic	5,688	36.41
			Proficient	5,033	32.22
			Advanced	2,774	17.76
			Proficient + Advanced	7,807	49.97
			Total	15,622	100.00
		Yes	Below Basic	2	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	2	100.00
	Geometry	No	Below Basic	503	14.07
			Basic	1,470	41.12
			Proficient	1,110	31.05
			Advanced	492	13.76
			Proficient + Advanced	1,602	44.81
			Total	3,575	100.00
		Yes	Below Basic	--	--
			Basic	--	--
			Proficient	1	100.00
			Advanced	--	--
			Proficient + Advanced	1	100.00
			Total	1	100.00
	Biology	No	Below Basic	10,132	16.78
			Basic	26,425	43.76
			Proficient	14,788	24.49
			Advanced	9,045	14.98
			Proficient + Advanced	23,833	39.47
			Total	60,390	100.00
		Yes	Below Basic	9	45.00
			Basic	9	45.00

Test Period	Content Area	Migrant	Achievement Level	Freq.	%
Spring 2019	Biology	Yes	Proficient	2	10.00
			Advanced	--	--
			Proficient + Advanced	2	10.00
			Total	20	100.00
	Physical Science	No	Below Basic	438	19.17
			Basic	997	43.63
			Proficient	638	27.92
			Advanced	212	9.28
			Proficient + Advanced	850	37.20
			Total	2,285	100.00
		Yes	Below Basic	1	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00

Table G.7. Achievement-Level Distributions—Free and Reduced Lunch, Fall 2018

Test Period	Content Area	FRL	Achievement Level	Freq.	%
Fall 2018	English I	No	Below Basic	2	4.35
			Basic	9	19.57
			Proficient	25	54.35
			Advanced	10	21.74
			Proficient + Advanced	35	76.09
			Total	46	100.00
		Yes	Below Basic	26	22.03
			Basic	41	34.75
			Proficient	36	30.51
			Advanced	15	12.71
			Proficient + Advanced	51	43.22
			Total	118	100.00
	English II	No	Below Basic	261	22.21
			Basic	258	21.96
			Proficient	541	46.04
			Advanced	115	9.79
			Proficient + Advanced	656	55.83
			Total	1,175	100.00
		Yes	Below Basic	538	39.39
			Basic	460	33.67
			Proficient	337	24.67
			Advanced	31	2.27
			Proficient + Advanced	368	26.94
			Total	1,366	100.00
	Algebra I	No	Below Basic	587	20.80
			Basic	682	24.17
			Proficient	609	21.58
			Advanced	944	33.45
			Proficient + Advanced	1,553	55.03
			Total	2,822	100.00
		Yes	Below Basic	1,218	48.94
			Basic	792	31.82
			Proficient	285	11.45
			Advanced	194	7.79
			Proficient + Advanced	479	19.24
			Total	2,489	100.00
	Algebra II	No	Below Basic	90	23.94
			Basic	85	22.61
			Proficient	134	35.64
			Advanced	67	17.82
			Proficient + Advanced	201	53.46
			Total	376	100.00
		Yes	Below Basic	74	43.53
			Basic	41	24.12
			Proficient	43	25.29
			Advanced	12	7.06

Test Period	Content Area	FRL	Achievement Level	Freq.	%
Fall 2018	Algebra II	Yes	Proficient + Advanced	55	32.35
			Total	170	100.00
	Geometry	No	Below Basic	1	1.11
			Basic	9	10.00
			Proficient	29	32.22
			Advanced	51	56.67
			Proficient + Advanced	80	88.89
			Total	90	100.00
		Yes	Below Basic	9	18.75
			Basic	13	27.08
			Proficient	11	22.92
			Advanced	15	31.25
			Proficient + Advanced	26	54.17
			Total	48	100.00
	Biology	No	Below Basic	255	19.54
			Basic	459	35.17
			Proficient	328	25.13
			Advanced	263	20.15
			Proficient + Advanced	591	45.29
			Total	1,305	100.00
		Yes	Below Basic	585	48.75
			Basic	474	39.50
			Proficient	103	8.58
			Advanced	38	3.17
			Proficient + Advanced	141	11.75
			Total	1,200	100.00
	Physical Science	No	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Yes	Below Basic	18	46.15
			Basic	18	46.15
			Proficient	3	7.69
			Advanced	--	--
			Proficient + Advanced	3	7.69
			Total	39	100.00

Table G.8. Achievement-Level Distributions—Free and Reduced Lunch, Spring 2019

Test Period	Content Area	FRL	Achievement Level	Freq.	%
Spring 2019	English I	No	Below Basic	244	4.83
			Basic	1,152	22.79
			Proficient	2,315	45.81
			Advanced	1,343	26.57
			Proficient + Advanced	3,658	72.38
			Total	5,054	100.00
		Yes	Below Basic	969	16.03
			Basic	2,069	34.22
			Proficient	2,206	36.49
			Advanced	802	13.26
			Proficient + Advanced	3,008	49.75
			Total	6,046	100.00
	English II	No	Below Basic	2,132	6.09
			Basic	8,075	23.08
			Proficient	19,446	55.58
			Advanced	5,336	15.25
			Proficient + Advanced	24,782	70.83
			Total	34,989	100.00
		Yes	Below Basic	4,931	18.59
			Basic	9,644	36.36
			Proficient	10,482	39.52
			Advanced	1,468	5.53
			Proficient + Advanced	11,950	45.05
			Total	26,525	100.00
	Algebra I	No	Below Basic	4,771	14.45
			Basic	9,102	27.57
			Proficient	8,965	27.15
			Advanced	10,178	30.83
			Proficient + Advanced	19,143	57.98
			Total	33,016	100.00
		Yes	Below Basic	9,198	34.42
			Basic	9,242	34.59
			Proficient	5,209	19.49
			Advanced	3,071	11.49
			Proficient + Advanced	8,280	30.99
			Total	26,720	100.00
	Algebra II	No	Below Basic	1,211	10.58
			Basic	3,840	33.55
			Proficient	3,978	34.76
			Advanced	2,415	21.10
			Proficient + Advanced	6,393	55.86
			Total	11,444	100.00
		Yes	Below Basic	918	21.96
			Basic	1,848	44.21
			Proficient	1,055	25.24
			Advanced	359	8.59
			Proficient + Advanced	1,414	33.83
			Total	4,180	100.00
	Geometry	No	Below Basic	252	11.23
			Basic	842	37.52
			Proficient	744	33.16

Test Period	Content Area	FRL	Achievement Level	Freq.	%
Spring 2019	Geometry	No	Advanced	406	18.09
			Proficient + Advanced	1,150	51.25
			Total	2,244	100.00
		Yes	Below Basic	251	18.84
			Basic	628	47.15
			Proficient	367	27.55
			Advanced	86	6.46
			Proficient + Advanced	453	34.01
			Total	1,332	100.00
	Biology	No	Below Basic	3,040	8.86
			Basic	14,016	40.85
			Proficient	10,161	29.61
			Advanced	7,094	20.68
			Proficient + Advanced	17,255	50.29
			Total	34,311	100.00
		Yes	Below Basic	7,101	27.21
			Basic	12,418	47.58
			Proficient	4,629	17.74
			Advanced	1,951	7.48
			Proficient + Advanced	6,580	25.21
			Total	26,099	100.00
	Physical Science	No	Below Basic	161	12.46
			Basic	544	42.11
			Proficient	421	32.59
			Advanced	166	12.85
			Proficient + Advanced	587	45.43
			Total	1,292	100.00
		Yes	Below Basic	278	27.97
			Basic	453	45.57
			Proficient	217	21.83
			Advanced	46	4.63
			Proficient + Advanced	263	26.46
			Total	994	100.00

Table G.9. Achievement-Level Distributions—Limited English Proficient, Fall 2018

Test Period	Content Area	LEP	Achievement Level	Freq.	%
Fall 2018	English I	No	Below Basic	21	15.91
			Basic	34	25.76
			Proficient	55	41.67
			Advanced	22	16.67
			Proficient + Advanced	77	58.33
			Total	132	100.00
		Yes	Below Basic	7	21.88
			Basic	16	50.00
			Proficient	6	18.75
			Advanced	3	9.38
			Proficient + Advanced	9	28.13
			Total	32	100.00
	English II	No	Below Basic	719	30.58
			Basic	638	27.14
			Proficient	848	36.07
			Advanced	146	6.21
			Proficient + Advanced	994	42.28
			Total	2,351	100.00
		Yes	Below Basic	80	42.11
			Basic	80	42.11
			Proficient	30	15.79
			Advanced	--	--
			Proficient + Advanced	30	15.79
			Total	190	100.00
	Algebra I	No	Below Basic	1,672	32.96
			Basic	1,396	27.52
			Proficient	877	17.29
			Advanced	1,128	22.24
			Proficient + Advanced	2,005	39.52
			Total	5,073	100.00
		Yes	Below Basic	133	55.88
			Basic	78	32.77
			Proficient	17	7.14
			Advanced	10	4.20
			Proficient + Advanced	27	11.34
			Total	238	100.00
	Algebra II	No	Below Basic	157	29.46
			Basic	125	23.45
			Proficient	172	32.27
			Advanced	79	14.82
			Proficient + Advanced	251	47.09
			Total	533	100.00
		Yes	Below Basic	7	53.85
			Basic	1	7.69
			Proficient	5	38.46
			Advanced	--	--
			Proficient + Advanced	5	38.46
			Total	13	100.00
	Geometry	No	Below Basic	10	7.25
			Basic	22	15.94
			Proficient	40	28.99

Test Period	Content Area	LEP	Achievement Level	Freq.	%
Fall 2018	Geometry	No	Advanced	66	47.83
			Proficient + Advanced	106	76.81
			Total	138	100.00
		Yes	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
	Biology	No	Below Basic	776	32.43
			Basic	892	37.28
			Proficient	425	17.76
			Advanced	300	12.54
			Proficient + Advanced	725	30.30
			Total	2,393	100.00
		Yes	Below Basic	64	57.14
			Basic	41	36.61
			Proficient	6	5.36
			Advanced	1	0.89
			Proficient + Advanced	7	6.25
			Total	112	100.00
	Physical Science	No	Below Basic	15	42.86
			Basic	17	48.57
			Proficient	3	8.57
			Advanced	--	--
			Proficient + Advanced	3	8.57
			Total	35	100.00
		Yes	Below Basic	3	75.00
			Basic	1	25.00
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	4	100.00

Table G.10. Achievement-Level Distributions—Limited English Proficient, Spring 2019

Test Period	Content Area	LEP	Achievement Level	Freq.	%
Spring 2019	English I	No	Below Basic	1,088	10.17
			Basic	3,072	28.72
			Proficient	4,407	41.20
			Advanced	2,129	19.90
			Proficient + Advanced	6,536	61.11
			Total	10,696	100.00
		Yes	Below Basic	125	30.94
			Basic	149	36.88
			Proficient	114	28.22
			Advanced	16	3.96
			Proficient + Advanced	130	32.18
			Total	404	100.00
	English II	No	Below Basic	6,493	10.88
			Basic	16,911	28.34
			Proficient	29,482	49.40
			Advanced	6,789	11.38
			Proficient + Advanced	36,271	60.78
			Total	59,675	100.00
		Yes	Below Basic	570	31.00
			Basic	808	43.94
			Proficient	446	24.25
			Advanced	15	0.82
			Proficient + Advanced	461	25.07
			Total	1,839	100.00
	Algebra I	No	Below Basic	13,140	22.83
			Basic	17,608	30.60
			Proficient	13,797	23.97
			Advanced	13,004	22.60
			Proficient + Advanced	26,801	46.57
			Total	57,549	100.00
		Yes	Below Basic	829	37.91
			Basic	736	33.65
			Proficient	377	17.24
			Advanced	245	11.20
			Proficient + Advanced	622	28.44
			Total	2,187	100.00
	Algebra II	No	Below Basic	2,059	13.40
			Basic	5,596	36.41
			Proficient	4,964	32.30
			Advanced	2,750	17.89
			Proficient + Advanced	7,714	50.19
			Total	15,369	100.00
		Yes	Below Basic	70	27.45
			Basic	92	36.08
			Proficient	69	27.06
			Advanced	24	9.41
			Proficient + Advanced	93	36.47
			Total	255	100.00
	Geometry	No	Below Basic	488	13.96
			Basic	1,429	40.88
			Proficient	1,090	31.18

Test Period	Content Area	LEP	Achievement Level	Freq.	%
Spring 2019	Geometry	No	Advanced	489	13.99
			Proficient + Advanced	1,579	45.17
			Total	3,496	100.00
		Yes	Below Basic	15	18.75
			Basic	41	51.25
			Proficient	21	26.25
			Advanced	3	3.75
			Proficient + Advanced	24	30.00
			Total	80	100.00
	Biology	No	Below Basic	9,287	15.88
			Basic	25,618	43.80
			Proficient	14,605	24.97
			Advanced	8,973	15.34
			Proficient + Advanced	23,578	40.32
			Total	58,483	100.00
		Yes	Below Basic	854	44.32
			Basic	816	42.35
			Proficient	185	9.60
			Advanced	72	3.74
			Proficient + Advanced	257	13.34
			Total	1,927	100.00
	Physical Science	No	Below Basic	433	19.10
			Basic	987	43.54
			Proficient	636	28.05
			Advanced	211	9.31
			Proficient + Advanced	847	37.36
			Total	2,267	100.00
		Yes	Below Basic	6	31.58
			Basic	10	52.63
			Proficient	2	10.53
			Advanced	1	5.26
			Proficient + Advanced	3	15.79
			Total	19	100.00

Table G.11. Achievement-Level Distributions—Title I, Fall 2018

Test Period	Content Area	Title I	Achievement Level	Freq.	%
Fall 2018	English I	No	Below Basic	8	80.00
			Basic	2	20.00
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	10	100.00
		Yes	Below Basic	20	12.99
			Basic	48	31.17
			Proficient	61	39.61
			Advanced	25	16.23
			Proficient + Advanced	86	55.84
			Total	154	100.00
	English II	No	Below Basic	625	30.96
			Basic	549	27.19
			Proficient	707	35.02
			Advanced	138	6.84
			Proficient + Advanced	845	41.85
			Total	2,019	100.00
		Yes	Below Basic	174	33.33
			Basic	169	32.38
			Proficient	171	32.76
			Advanced	8	1.53
			Proficient + Advanced	179	34.29
			Total	522	100.00
	Algebra I	No	Below Basic	1,430	31.30
			Basic	1,225	26.81
			Proficient	812	17.77
			Advanced	1,102	24.12
			Proficient + Advanced	1,914	41.89
			Total	4,569	100.00
		Yes	Below Basic	375	50.54
			Basic	249	33.56
			Proficient	82	11.05
			Advanced	36	4.85
			Proficient + Advanced	118	15.90
			Total	742	100.00
	Algebra II	No	Below Basic	122	25.47
			Basic	112	23.38
			Proficient	166	34.66
			Advanced	79	16.49
			Proficient + Advanced	245	51.15
			Total	479	100.00
		Yes	Below Basic	42	62.69
			Basic	14	20.90
			Proficient	11	16.42
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	67	100.00
	Geometry	No	Below Basic	1	0.88
			Basic	10	8.85
			Proficient	37	32.74

Test Period	Content Area	Title I	Achievement Level	Freq.	%
Fall 2018	Geometry	No	Advanced	65	57.52
			Proficient + Advanced	102	90.27
			Total	113	100.00
		Yes	Below Basic	9	36.00
			Basic	12	48.00
			Proficient	3	12.00
			Advanced	1	4.00
			Proficient + Advanced	4	16.00
			Total	25	100.00
	Biology	No	Below Basic	639	30.04
			Basic	778	36.58
			Proficient	412	19.37
			Advanced	298	14.01
			Proficient + Advanced	710	33.38
			Total	2,127	100.00
		Yes	Below Basic	201	53.17
			Basic	155	41.01
			Proficient	19	5.03
			Advanced	3	0.79
			Proficient + Advanced	22	5.82
			Total	378	100.00
	Physical Science	No	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--
		Yes	Below Basic	18	46.15
			Basic	18	46.15
			Proficient	3	7.69
			Advanced	--	--
			Proficient + Advanced	3	7.69
			Total	39	100.00

Table G.12. Achievement-Level Distributions—Title I, Spring 2019

Test Period	Content Area	Title I	Achievement Level	Freq.	%
Spring 2019	English I	No	Below Basic	761	8.32
			Basic	2,557	27.95
			Proficient	3,898	42.61
			Advanced	1,933	21.13
			Proficient + Advanced	5,831	63.73
			Total	9,149	100.00
		Yes	Below Basic	452	23.17
			Basic	664	34.03
			Proficient	623	31.93
			Advanced	212	10.87
			Proficient + Advanced	835	42.80
			Total	1,951	100.00
	English II	No	Below Basic	5,772	10.27
			Basic	15,636	27.83
			Proficient	28,171	50.14
			Advanced	6,608	11.76
			Proficient + Advanced	34,779	61.90
			Total	56,187	100.00
		Yes	Below Basic	1,291	24.24
			Basic	2,083	39.10
			Proficient	1,757	32.98
			Advanced	196	3.68
			Proficient + Advanced	1,953	36.66
			Total	5,327	100.00
	Algebra I	No	Below Basic	11,405	21.30
			Basic	16,351	30.54
			Proficient	13,150	24.56
			Advanced	12,637	23.60
			Proficient + Advanced	25,787	48.16
			Total	53,543	100.00
		Yes	Below Basic	2,564	41.40
			Basic	1,993	32.18
			Proficient	1,024	16.53
			Advanced	612	9.88
			Proficient + Advanced	1,636	26.42
			Total	6,193	100.00
	Algebra II	No	Below Basic	1,933	12.96
			Basic	5,392	36.15
			Proficient	4,860	32.58
			Advanced	2,732	18.31
			Proficient + Advanced	7,592	50.89
			Total	14,917	100.00
		Yes	Below Basic	196	27.72
			Basic	296	41.87
			Proficient	173	24.47
			Advanced	42	5.94
			Proficient + Advanced	215	30.41
			Total	707	100.00
	Geometry	No	Below Basic	470	13.93
			Basic	1,374	40.72
			Proficient	1,053	31.21

Test Period	Content Area	Title I	Achievement Level	Freq.	%
Spring 2019	Geometry	No	Advanced	477	14.14
			Proficient + Advanced	1,530	45.35
			Total	3,374	100.00
		Yes	Below Basic	33	16.34
			Basic	96	47.52
			Proficient	58	28.71
			Advanced	15	7.43
			Proficient + Advanced	73	36.14
			Total	202	100.00
	Biology	No	Below Basic	8,036	14.73
			Basic	23,802	43.62
			Proficient	13,978	25.61
			Advanced	8,756	16.04
			Proficient + Advanced	22,734	41.66
			Total	54,572	100.00
		Yes	Below Basic	2,105	36.06
			Basic	2,632	45.08
			Proficient	812	13.91
			Advanced	289	4.95
			Proficient + Advanced	1,101	18.86
			Total	5,838	100.00
	Physical Science	No	Below Basic	412	19.00
			Basic	939	43.31
			Proficient	608	28.04
			Advanced	209	9.64
			Proficient + Advanced	817	37.68
			Total	2,168	100.00
		Yes	Below Basic	27	22.88
			Basic	58	49.15
			Proficient	30	25.42
			Advanced	3	2.54
			Proficient + Advanced	33	27.97
			Total	118	100.00

Table G.13. Achievement-Level Distributions—Individualized Education Program, Fall 2018

Test Period	Content Area	IEP	Achievement Level	Freq.	%
Fall 2018	English I	No	Below Basic	25	16.03
			Basic	45	28.85
			Proficient	61	39.10
			Advanced	25	16.03
			Proficient + Advanced	86	55.13
			Total	156	100.00
		Yes	Below Basic	3	37.50
			Basic	5	62.50
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	8	100.00
	English II	No	Below Basic	602	26.88
			Basic	634	28.30
			Proficient	858	38.30
			Advanced	146	6.52
			Proficient + Advanced	1,004	44.82
			Total	2,240	100.00
		Yes	Below Basic	197	65.45
			Basic	84	27.91
			Proficient	20	6.64
			Advanced	--	--
			Proficient + Advanced	20	6.64
			Total	301	100.00
	Algebra I	No	Below Basic	1,413	30.02
			Basic	1,327	28.19
			Proficient	861	18.29
			Advanced	1,106	23.50
			Proficient + Advanced	1,967	41.79
			Total	4,707	100.00
		Yes	Below Basic	392	64.90
			Basic	147	24.34
			Proficient	33	5.46
			Advanced	32	5.30
			Proficient + Advanced	65	10.76
			Total	604	100.00
	Algebra II	No	Below Basic	160	29.80
			Basic	124	23.09
			Proficient	175	32.59
			Advanced	78	14.53
			Proficient + Advanced	253	47.11
			Total	537	100.00
		Yes	Below Basic	4	44.44
			Basic	2	22.22
			Proficient	2	22.22
			Advanced	1	11.11
			Proficient + Advanced	3	33.33
			Total	9	100.00
	Geometry	No	Below Basic	10	7.35
			Basic	21	15.44
			Proficient	40	29.41

Test Period	Content Area	IEP	Achievement Level	Freq.	%
Fall 2018	Geometry	No	Advanced	65	47.79
			Proficient + Advanced	105	77.21
			Total	136	100.00
		Yes	Below Basic	--	--
			Basic	1	50.00
			Proficient	--	--
			Advanced	1	50.00
			Proficient + Advanced	1	50.00
			Total	2	100.00
	Biology	No	Below Basic	656	29.80
			Basic	838	38.07
			Proficient	414	18.81
			Advanced	293	13.31
			Proficient + Advanced	707	32.12
			Total	2,201	100.00
		Yes	Below Basic	184	60.53
			Basic	95	31.25
			Proficient	17	5.59
			Advanced	8	2.63
			Proficient + Advanced	25	8.22
			Total	304	100.00
	Physical Science	No	Below Basic	17	44.74
			Basic	18	47.37
			Proficient	3	7.89
			Advanced	--	--
			Proficient + Advanced	3	7.89
			Total	38	100.00
		Yes	Below Basic	1	100.00
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	1	100.00

Table G.14. Achievement-Level Distributions—Individualized Education Program, Spring 2019

Test Period	Content Area	IEP	Achievement Level	Freq.	%
Spring 2019	English I	No	Below Basic	826	8.19
			Basic	2,759	27.35
			Proficient	4,379	43.41
			Advanced	2,123	21.05
			Proficient + Advanced	6,502	64.46
			Total	10,087	100.00
		Yes	Below Basic	387	38.20
			Basic	462	45.61
			Proficient	142	14.02
			Advanced	22	2.17
			Proficient + Advanced	164	16.19
			Total	1,013	100.00
	English II	No	Below Basic	4,400	8.02
			Basic	14,985	27.30
			Proficient	28,776	52.42
			Advanced	6,733	12.27
			Proficient + Advanced	35,509	64.69
			Total	54,894	100.00
		Yes	Below Basic	2,663	40.23
			Basic	2,734	41.30
			Proficient	1,152	17.40
			Advanced	71	1.07
			Proficient + Advanced	1,223	18.47
			Total	6,620	100.00
	Algebra I	No	Below Basic	9,981	18.75
			Basic	16,703	31.38
			Proficient	13,596	25.55
			Advanced	12,940	24.31
			Proficient + Advanced	26,536	49.86
			Total	53,220	100.00
		Yes	Below Basic	3,988	61.20
			Basic	1,641	25.18
			Proficient	578	8.87
			Advanced	309	4.74
			Proficient + Advanced	887	13.61
			Total	6,516	100.00
	Algebra II	No	Below Basic	2,053	13.33
			Basic	5,611	36.43
			Proficient	4,989	32.39
			Advanced	2,750	17.85
			Proficient + Advanced	7,739	50.24
			Total	15,403	100.00
		Yes	Below Basic	76	34.39
			Basic	77	34.84
			Proficient	44	19.91
			Advanced	24	10.86
			Proficient + Advanced	68	30.77
			Total	221	100.00
	Geometry	No	Below Basic	449	13.11
			Basic	1,405	41.01
			Proficient	1,086	31.70

Test Period	Content Area	IEP	Achievement Level	Freq.	%
Spring 2019	Geometry	No	Advanced	486	14.19
			Proficient + Advanced	1,572	45.88
			Total	3,426	100.00
		Yes	Below Basic	54	36.00
			Basic	65	43.33
			Proficient	25	16.67
			Advanced	6	4.00
			Proficient + Advanced	31	20.67
			Total	150	100.00
	Biology	No	Below Basic	6,993	12.95
			Basic	23,897	44.26
			Proficient	14,253	26.40
			Advanced	8,849	16.39
			Proficient + Advanced	23,102	42.79
			Total	53,992	100.00
		Yes	Below Basic	3,148	49.05
			Basic	2,537	39.53
			Proficient	537	8.37
			Advanced	196	3.05
			Proficient + Advanced	733	11.42
			Total	6,418	100.00
	Physical Science	No	Below Basic	346	16.37
			Basic	925	43.78
			Proficient	631	29.86
			Advanced	211	9.99
			Proficient + Advanced	842	39.85
			Total	2,113	100.00
		Yes	Below Basic	93	53.76
			Basic	72	41.62
			Proficient	7	4.05
			Advanced	1	0.58
			Proficient + Advanced	8	4.62
			Total	173	100.00

Table G.15. Achievement-Level Distributions—Accommodations, Fall 2018

Test Period	Content Area	Accom.	Achievement Level	Freq.	%
Fall 2018	English I	No	Below Basic	22	15.17
			Basic	41	28.28
			Proficient	57	39.31
			Advanced	25	17.24
			Proficient + Advanced	82	56.55
			Total	145	100.00
		Yes	Below Basic	6	31.58
			Basic	9	47.37
			Proficient	4	21.05
			Advanced	--	--
			Proficient + Advanced	4	21.05
			Total	19	100.00
	English II	No	Below Basic	649	28.55
			Basic	635	27.94
			Proficient	844	37.13
			Advanced	145	6.38
			Proficient + Advanced	989	43.51
			Total	2,273	100.00
		Yes	Below Basic	150	55.97
			Basic	83	30.97
			Proficient	34	12.69
			Advanced	1	0.37
			Proficient + Advanced	35	13.06
			Total	268	100.00
	Algebra I	No	Below Basic	1,531	31.60
			Basic	1,353	27.93
			Proficient	852	17.59
			Advanced	1,109	22.89
			Proficient + Advanced	1,961	40.47
			Total	4,845	100.00
		Yes	Below Basic	274	58.80
			Basic	121	25.97
			Proficient	42	9.01
			Advanced	29	6.22
			Proficient + Advanced	71	15.24
			Total	466	100.00
	Algebra II	No	Below Basic	162	30.00
			Basic	125	23.15
			Proficient	175	32.41
			Advanced	78	14.44
			Proficient + Advanced	253	46.85
			Total	540	100.00
		Yes	Below Basic	2	33.33
			Basic	1	16.67
			Proficient	2	33.33
			Advanced	1	16.67
			Proficient + Advanced	3	50.00
			Total	6	100.00
	Geometry	No	Below Basic	10	7.30
			Basic	22	16.06
			Proficient	40	29.20

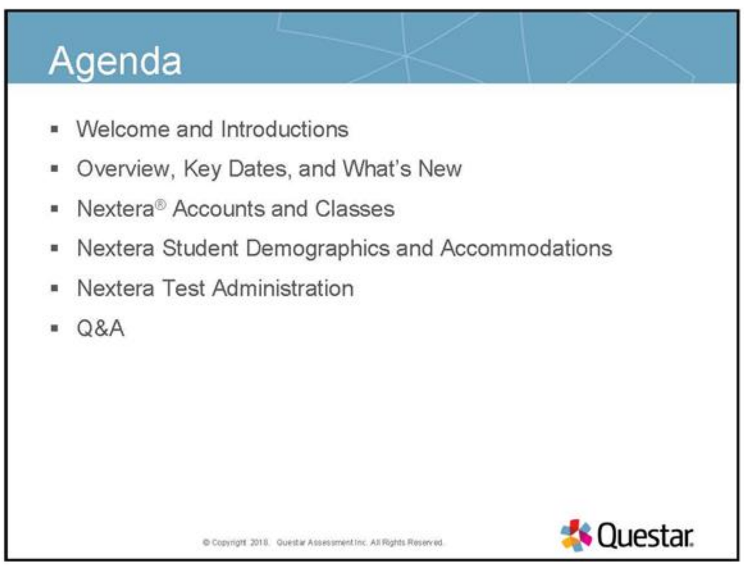
Test Period	Content Area	Accom.	Achievement Level	Freq.	%
Fall 2018	Geometry	No	Advanced	65	47.45
			Proficient + Advanced	105	76.64
			Total	137	100.00
		Yes	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	1	100.00
			Proficient + Advanced	1	100.00
			Total	1	100.00
	Biology	No	Below Basic	698	31.01
			Basic	839	37.27
			Proficient	420	18.66
			Advanced	294	13.06
			Proficient + Advanced	714	31.72
			Total	2,251	100.00
		Yes	Below Basic	142	55.91
			Basic	94	37.01
			Proficient	11	4.33
			Advanced	7	2.76
			Proficient + Advanced	18	7.09
			Total	254	100.00
	Physical Science	No	Below Basic	18	46.15
			Basic	18	46.15
			Proficient	3	7.69
			Advanced	--	--
			Proficient + Advanced	3	7.69
			Total	39	100.00
		Yes	Below Basic	--	--
			Basic	--	--
			Proficient	--	--
			Advanced	--	--
			Proficient + Advanced	--	--
			Total	--	--

Table G.16. Achievement-Level Distributions—Accommodations, Spring 2019

Test Period	Content Area	Accom.	Achievement Level	Freq.	%
Spring 2019	English I	No	Below Basic	890	9.09
			Basic	2,701	27.57
			Proficient	4,182	42.69
			Advanced	2,023	20.65
			Proficient + Advanced	6,205	63.34
			Total	9,796	100.00
		Yes	Below Basic	323	24.77
			Basic	520	39.88
			Proficient	339	26.00
			Advanced	122	9.36
			Proficient + Advanced	461	35.35
			Total	1,304	100.00
	English II	No	Below Basic	4,603	8.72
			Basic	14,494	27.45
			Proficient	27,315	51.73
			Advanced	6,387	12.10
			Proficient + Advanced	33,702	63.83
			Total	52,799	100.00
		Yes	Below Basic	2,460	28.23
			Basic	3,225	37.01
			Proficient	2,613	29.98
			Advanced	417	4.78
			Proficient + Advanced	3,030	34.77
			Total	8,715	100.00
	Algebra I	No	Below Basic	10,009	19.37
			Basic	16,028	31.02
			Proficient	13,076	25.31
			Advanced	12,558	24.30
			Proficient + Advanced	25,634	49.61
			Total	51,671	100.00
		Yes	Below Basic	3,960	49.10
			Basic	2,316	28.72
			Proficient	1,098	13.61
			Advanced	691	8.57
			Proficient + Advanced	1,789	22.18
			Total	8,065	100.00
	Algebra II	No	Below Basic	1,976	13.42
			Basic	5,294	35.95
			Proficient	4,783	32.48
			Advanced	2,673	18.15
			Proficient + Advanced	7,456	50.63
			Total	14,726	100.00
		Yes	Below Basic	153	17.04
			Basic	394	43.88
			Proficient	250	27.84
			Advanced	101	11.25
			Proficient + Advanced	351	39.09
			Total	898	100.00
	Geometry	No	Below Basic	450	13.17
			Basic	1,401	41.01
			Proficient	1,080	31.62

Test Period	Content Area	Accom.	Achievement Level	Freq.	%
Spring 2019	Geometry	No	Advanced	485	14.20
			Proficient + Advanced	1,565	45.81
			Total	3,416	100.00
		Yes	Below Basic	53	33.13
			Basic	69	43.13
			Proficient	31	19.38
			Advanced	7	4.38
			Proficient + Advanced	38	23.75
			Total	160	100.00
	Biology	No	Below Basic	6,941	13.46
			Basic	22,643	43.90
			Proficient	13,549	26.27
			Advanced	8,444	16.37
			Proficient + Advanced	21,993	42.64
			Total	51,577	100.00
		Yes	Below Basic	3,200	36.23
			Basic	3,791	42.92
			Proficient	1,241	14.05
			Advanced	601	6.80
			Proficient + Advanced	1,842	20.85
			Total	8,833	100.00
	Physical Science	No	Below Basic	356	16.77
			Basic	930	43.81
			Proficient	628	29.58
			Advanced	209	9.84
			Proficient + Advanced	837	39.43
			Total	2,123	100.00
		Yes	Below Basic	83	50.92
			Basic	67	41.10
			Proficient	10	6.13
			Advanced	3	1.84
			Proficient + Advanced	13	7.98
			Total	163	100.00

Appendix H: Test Coordinator Training PowerPoints



Missouri End-of-Course Assessments Overview

General Information

- New EOC assessments aligned to the new standards have been built for all Mathematics and English Language Arts courses.
- Biology EOC testing in 2017-2018 will be a Stand Alone Field Test. Government and American History EOC assessments will remain the same throughout 2017-2018.

Questar

Tutorials

- New Tutorials are available!
  - Generic with all item types
  - Math tutorial with math item types
  - ELA tutorial with ELA item types
  - Science tutorial with science item types
- Offer an opportunity for students to become familiar with the item types, tools and format they will experience during testing.
- Administrators are encouraged to allow students plenty of time to work with the tutorials to become familiar with all the new item types and testing platform.

Questar

Nextera

Nextera System Overview

Nextera, is made up of two components that provide a full-service assessment solution.

- Online Test Administration System (Nextera Admin):
  - Student and testing management tools
  - Multi-tiered, role-based system
- Test Delivery System
  - The Secure Browser keeps students focused on their test
  - Test content is downloaded to the student's device to ensure uninterrupted testing for students

Questar

Nextera Admin

- Secure, web-based administration system provides access to all users with no additional download.
  - Login
  - Home
  - Students
  - Classes
  - Test Administrations
  - Accounts
  - Reports
  - Help

Questar

Nextera Test Delivery System

- Installed on each device
- Allows the test to be presented securely on the device
- Employs an HTML5 framework
  - No Java dependencies
- Provides confidence in saving student responses
  - Test content cached when student logs in
    - Student response/interaction continuously sent to Questar
    - Responses stored/encrypted locally on computer/device in case of network loss

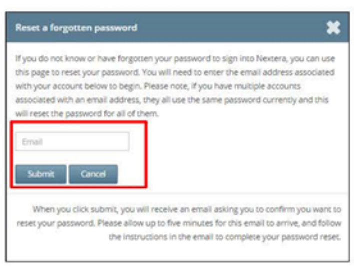
Questar

Sign in

If you forget or lose your password:

1. Click the "**Forgot your password?**" link.

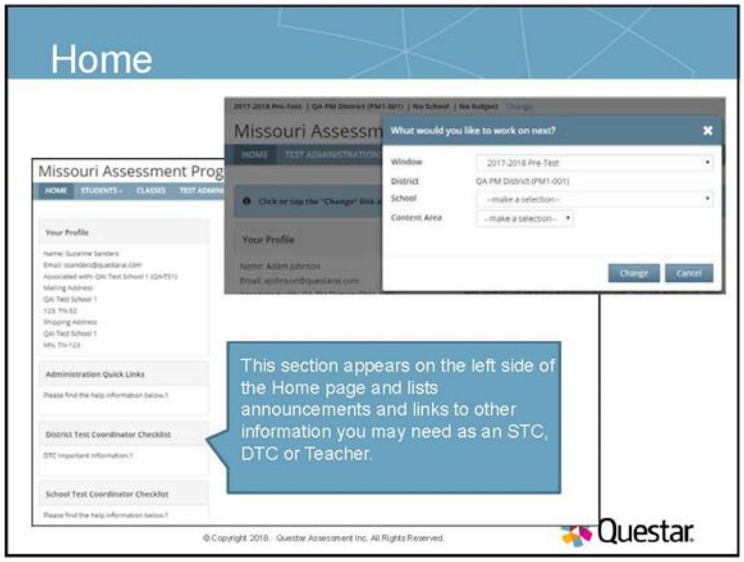
2. From the Forgotten Password screen, enter email address and click "**Submit**".

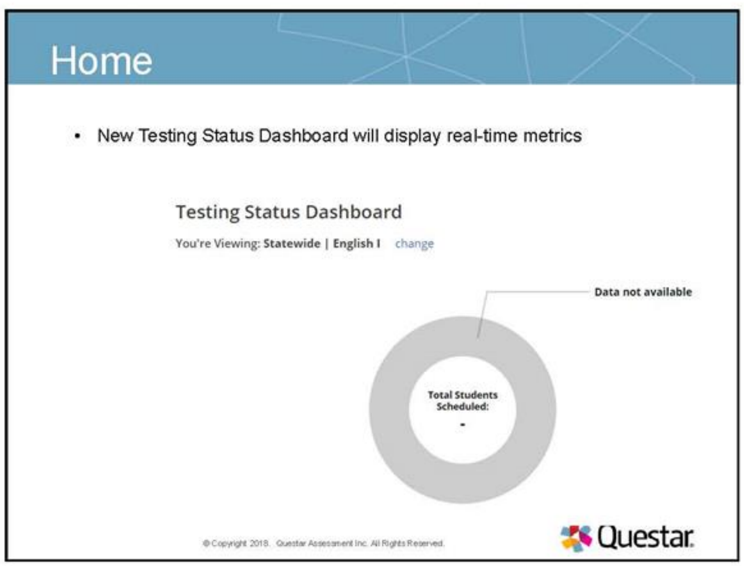


3. Receive email to confirm you want to reset your password.

4. Please allow up to five minutes for this email to arrive, and follow the instructions in the email to complete your password reset.

Questar

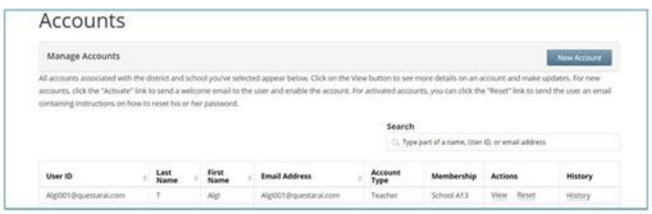

Home
Missouri Assessment Prog
Missouri Assessm
What would you like to work on next?
Window
2017-2018 Pre-Test
District
QA PM District (PM1-001)
School
--make a selection--
Content Area
--make a selection--
Change
Cancel
Your Profile
Administration Quick Links
District Test Coordinator Checklist
School Test Coordinator Checklist
This section appears on the left side of the Home page and lists announcements and links to other information you may need as an STC, DTC or Teacher.
© Copyright 2018. Questar Assessment Inc. All Rights Reserved.
Questar


Home
• New Testing Status Dashboard will display real-time metrics
Testing Status Dashboard
You're Viewing: Statewide | English I change
Data not available
Total Students Scheduled:
-
© Copyright 2018. Questar Assessment Inc. All Rights Reserved.
Questar

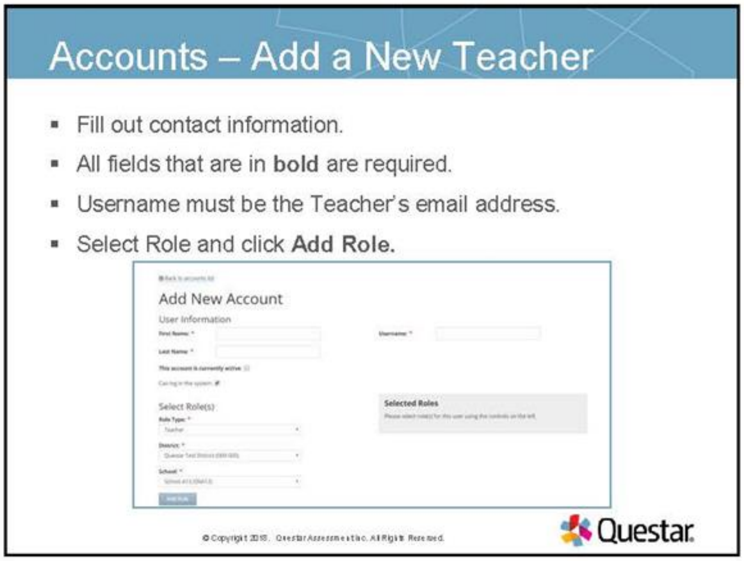
Nextera Admin Accounts and Classes

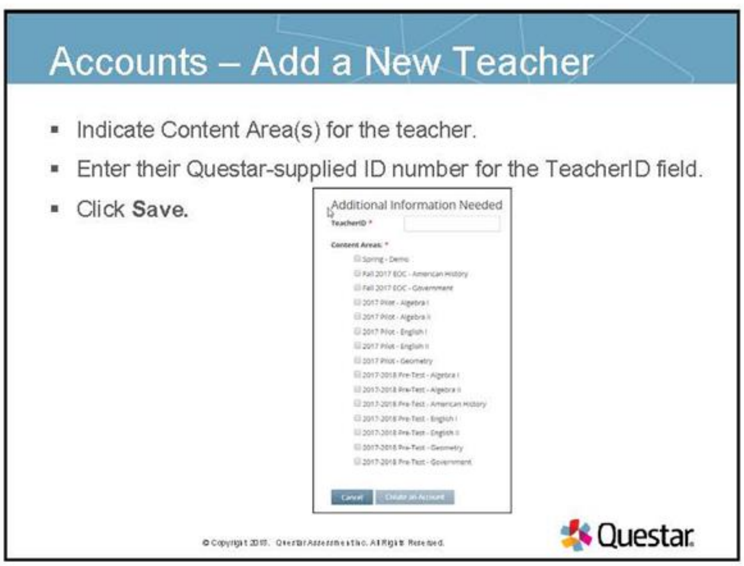
Accounts

- DTCs and STCs are responsible for managing profiles in the Nextera Admin site.
  - Select the Accounts menu, then select Accounts.
  - Validate the Teachers listed.
  - To add a new Teacher, click **New Account.**



Questar.


Accounts – Add a New Teacher
Fill out contact information.
All fields that are in **bold** are required.
Username must be the Teacher's email address.
Select Role and click **Add Role.**
Add New Account
User Information
Select Role(s)
Selected Roles
Questar


Accounts – Add a New Teacher
Indicate Content Area(s) for the teacher.
Enter their Questar-supplied ID number for the TeacherID field.
Click **Save.**
Additional Information Needed
Questar

Accounts

- To Edit an existing Teacher's account, click **View** then click **Edit** on the View Teacher screen.
- Click **Reset** to send a temporary password to a Teacher.
  - This new password will replace the temporary password and will become the password they should use for future logins.

Accounts

Manage Accounts — New Account

All accounts associated with the district and school you've selected appear below. Click on the View button to see more details on an account and make updates. For new accounts, click the "Activate" link to send a welcome email to the user and enable the account. For activated accounts, you can click the "Reset" link to send the user an email containing instructions on how to reset his or her password.

Search

Type part of a name, User ID, or email address

User ID	Last Name	First Name	Email Address	Account Type	Membership	Actions	History
Algi001@questarai.com	T	Algi	Algi001@questarai.com	Teacher	School A13	View Reset	History

Questar.

Accounts – District Testing Window

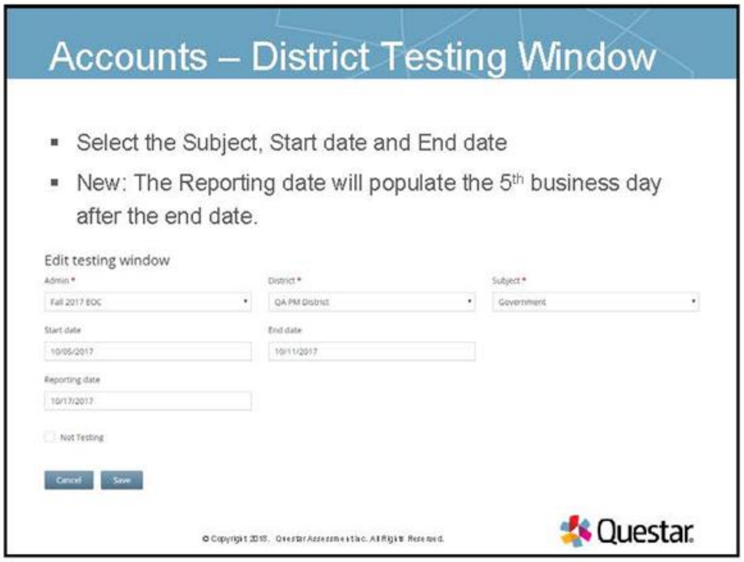
- Located under the Accounts tab, DTCs need to indicate a 7-Day testing window for each content area they plan to test.
- Click "Create Test Window"

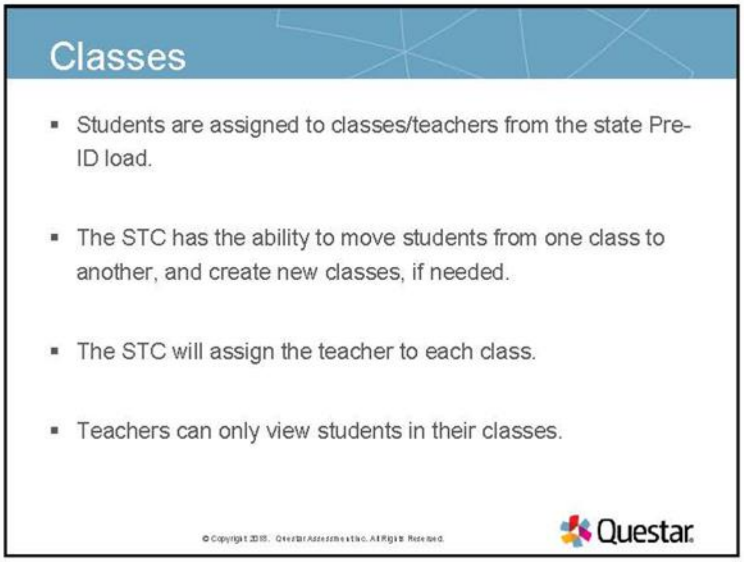
District Test Windows

Create Test Window

Admin Name	District	Subject Name	Start Date	End Date	
Fall 2017 EOC	QA PM District	Government	2017-10-05	2017-10-11	Edit
Fall 2017 EOC	QA PM District	American History	2017-10-09	2017-10-17	Edit
Fall 2017 EOC	QA PM District	Algebra I	2017-11-06	2017-11-14	Edit
Fall 2017 EOC	QA PM District	Algebra II	2017-11-14	2017-11-22	Edit
Fall 2017 EOC	QA PM District	English I	2017-11-13	2017-11-21	Edit
Fall 2017 EOC	QA PM District	Geometry	2017-11-13	2017-11-21	Edit
Fall 2017 EOC	QA PM District	English II	2017-11-13	2017-11-21	Edit
Fall 2017 EOC	QA PM District	Biology	2017-11-13	2017-11-21	Edit

Questar.


Accounts – District Testing Window
Select the Subject, Start date and End date
New: The Reporting date will populate the 5th business day after the end date.
Edit testing window
Admin *
Fall 2017 EOC
District *
QA PM District
Subject *
Government
Start date
10/05/2017
End date
10/11/2017
Reporting date
10/17/2017
Not Testing
Cancel
Save
Questar


Classes
Students are assigned to classes/teachers from the state Pre-ID load.
The STC has the ability to move students from one class to another, and create new classes, if needed.
The STC will assign the teacher to each class.
Teachers can only view students in their classes.
Questar

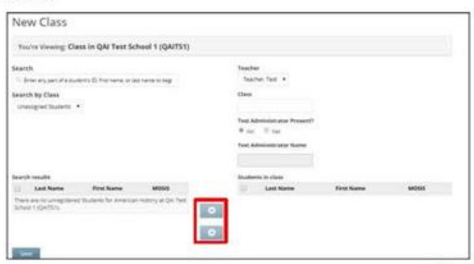
Classes – New Class

- To add a new class, from the Classes page, click **New Class.**
- Select the Teacher from drop-down.
- Name the Class.

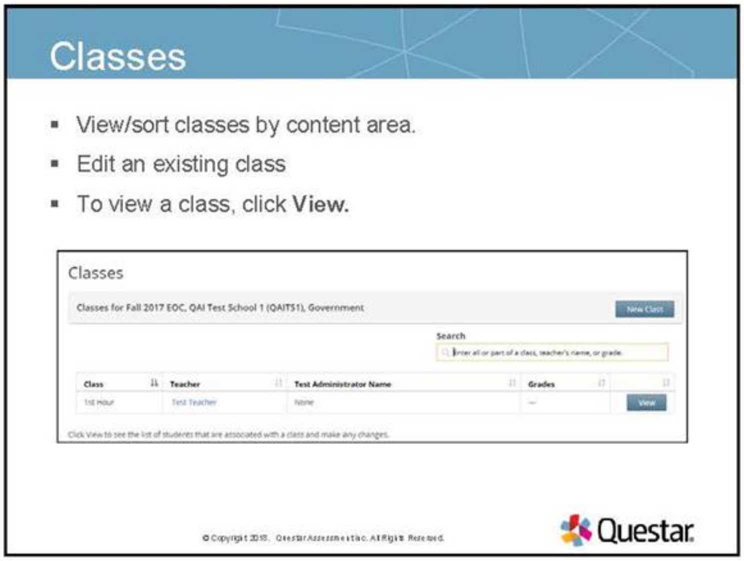
Questar

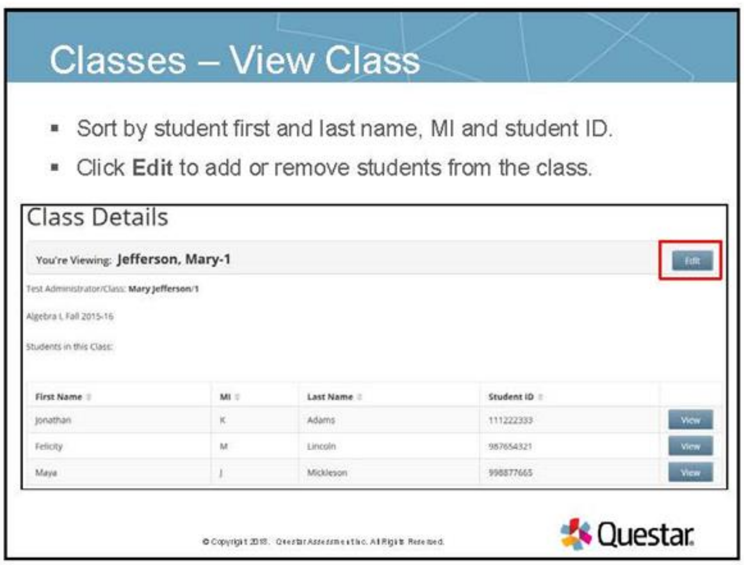
Classes – New Class continued

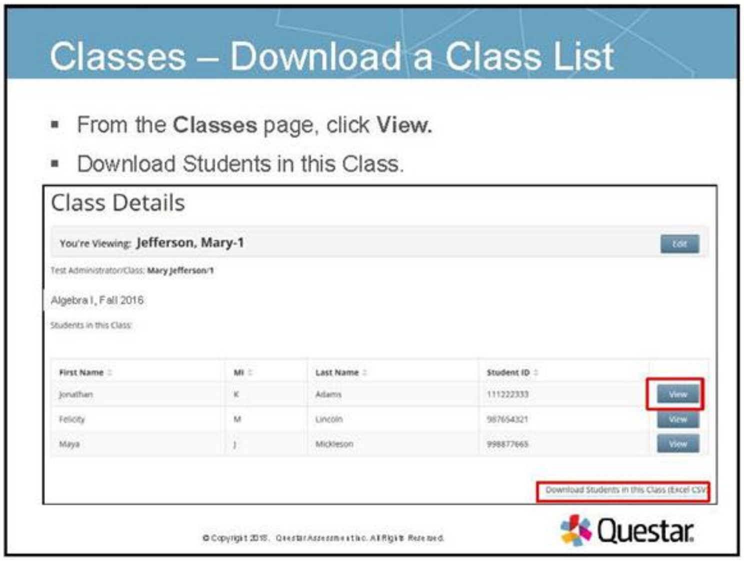
- Select from different grouping of students using the Class drop-down menu.
- Mark the students you want to add to the class using arrows.
- Click **Save.**



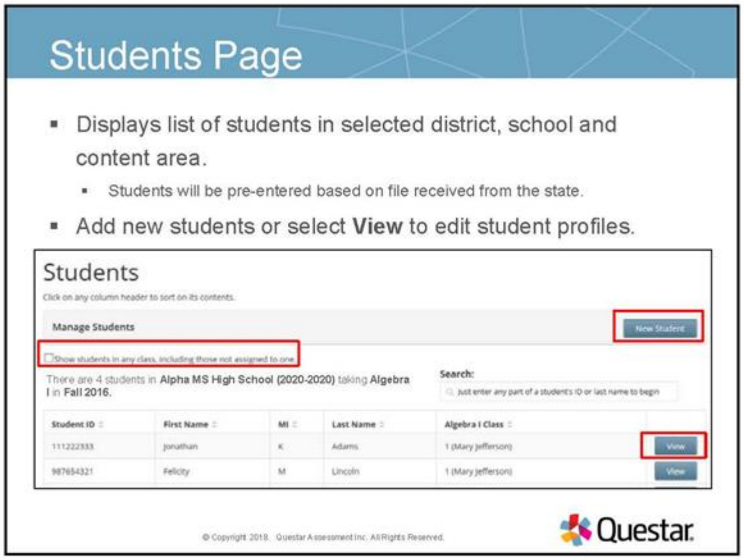
Questar

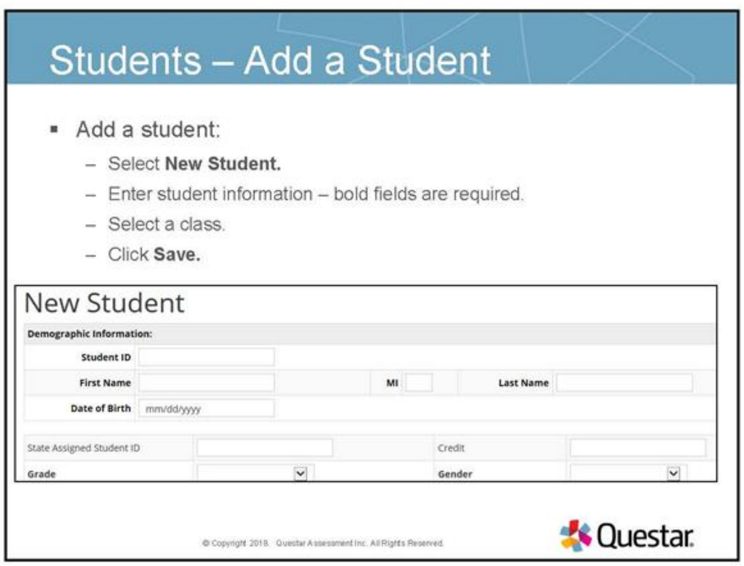

Classes
View/sort classes by content area.
Edit an existing class
To view a class, click View.
Classes
Classes for Fall 2017 EOC, QAI Test School 1 (QAITS1), Government
New Class
Search
Enter all or part of a class, teacher's name, or grade.
Class
Teacher
Test Administrator Name
Grades
1st Hour
Test Teacher
None
View
Click View to see the list of students that are associated with a class and make any changes.
Questar


Classes – View Class
Sort by student first and last name, MI and student ID.
Click Edit to add or remove students from the class.
Class Details
You're Viewing: Jefferson, Mary-1
Edit
Test Administrator/Class: Mary Jefferson/1
Algebra I, Fall 2015-16
Students in this Class:
First Name
MI
Last Name
Student ID
Jonathan
K
Adams
111222333
View
Felicity
M
Lincoln
987654321
View
Maya
J
Mickleson
998877665
View
Questar


Classes – Download a Class List
From the **Classes** page, click **View.**
Download Students in this Class.
Class Details
You're Viewing: Jefferson, Mary-1
Edit
Test Administrator/Class: Mary Jefferson/1
Algebra I, Fall 2016
Students in this Class:
First Name
MI
Last Name
Student ID
Jonathan
K
Adams
111222333
View
Felicity
M
Lincoln
987654321
View
Maya
J
Mickleson
998877665
View
Download Students in this Class (Excel CSV)
Questar

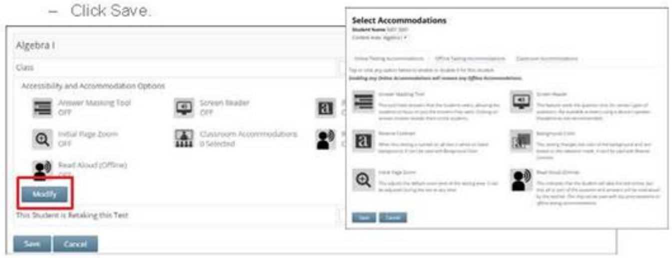

Nextera Admin
Students & Accommodations


Students Page
Displays list of students in selected district, school and content area.
Students will be pre-entered based on file received from the state.
Add new students or select **View** to edit student profiles.
Students
Click on any column header to sort on its contents.
Manage Students
New Student
Show students in any class, including those not assigned to one
There are 4 students in Alpha MS High School (2020-2020) taking Algebra I in Fall 2016.
Search:
Just enter any part of a student's ID or last name to begin
Student ID
First Name
MI
Last Name
Algebra I Class
111222333
Jonathan
K
Adams
1 (Mary Jefferson)
View
987654321
Felicity
M
Lincoln
1 (Mary Jefferson)
View
© Copyright 2018. Questar Assessment Inc. All Rights Reserved.
Questar


Students – Add a Student
Add a student:
– Select **New Student.**
– Enter student information – bold fields are required.
– Select a class.
– Click **Save.**
New Student
Demographic Information:
Student ID
First Name
MI
Last Name
Date of Birth
mm/dd/yyyy
State Assigned Student ID
Credit
Grade
Gender
© Copyright 2018. Questar Assessment Inc. All Rights Reserved.
Questar

Students – Edit a Student

- To edit a student:
  - Select **View,** then **Edit.**
  - Edit any fields that are not grayed out.
  - Set Accessibility and Accommodations for a student by clicking **Modify.**
  - Select the approved options.
  - Click Save.



Accommodations

- Accommodations and accessibility options can be set for each test a student is taking.

Accommodations Continued

- Online Accommodations highlights
  - Text-To-Speech
  - Read Aloud
- Offline Accommodations highlights
  - Print variations such as Paper, Large Print and Braille
- Classroom Accommodations highlights
  - Multi select tab to identify any classroom accommodations
  - Read aloud options, scribe. Etc.
- These accommodations must be indicated prior to testing.

Questar

NEW: Students – Multi Student Edit

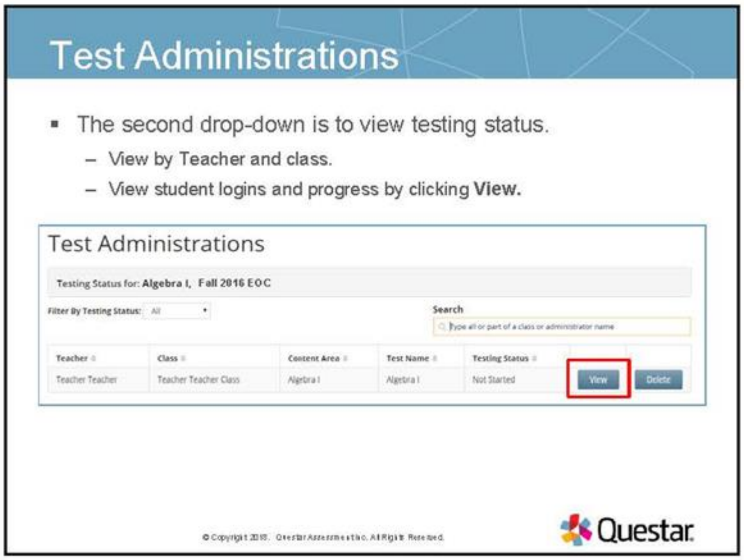
- Multi Student Edit:
  - Select **Multi Student Edit** from the Students Tab dropdown.
  - Select the desired tool or accommodation (it will turn blue)
  - Check boxes for students.
  - Click Save.

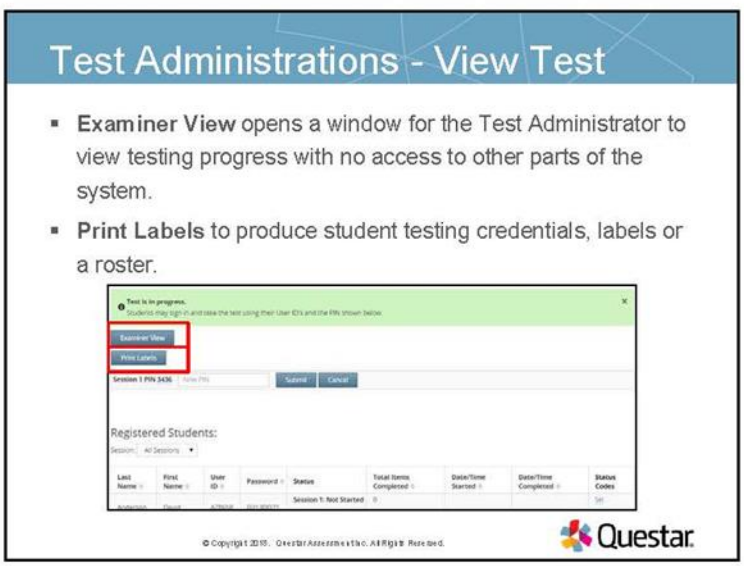
Questar

Nextera Admin
Test Administration

Important Dates

- The first drop-down on the Test Administrations menu references important dates throughout the academic year.


Test Administrations
The second drop-down is to view testing status.
– View by Teacher and class.
– View student logins and progress by clicking View.
Test Administrations
Testing Status for: Algebra I, Fall 2016 EOC
Filter By Testing Status: All
Search
Type all or part of a class or administrator name
Teacher
Class
Content Area
Test Name
Testing Status
Teacher Teacher
Teacher Teacher Class
Algebra I
Algebra I
Not Started
View
Delete
Questar


Test Administrations - View Test
Examiner View opens a window for the Test Administrator to view testing progress with no access to other parts of the system.
Print Labels to produce student testing credentials, labels or a roster.
Test is in progress.
Examiner View
Print Labels
Submit
Cancel
Registered Students:
Last Name
First Name
User ID
Password
Status
Total Items Completed
Date/Time Started
Date/Time Completed
Status Codes
Session 1: Not Started
Questar

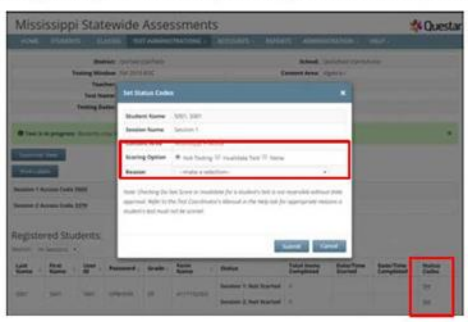
Student Invalidations

- If a student's test session needs to be invalidated, the test administrator should notify the DTC immediately following the invalid test session.
  1. The DTC will contact DESE and complete an irregularity report.
  2. Enter the invalidation into the Nextera Admin site by clicking set status Codes.

Status	Total Items Completed	Date/Time Started	Date/Time Completed	Status Codes
Session 1: Not Started	0			Set

Test Administrations – Status Codes

- Indicate *Not Testing* or *Invalidate Test* and then select the reason from the dropdown box.
- If invalidating an Algebra I student, ensure you invalidate both sessions (1&2).



NEW: Pre-Test Teacher Scoring

- Pre-Test Teacher Scoring is now done via the Admin site
  - All constructed response item types.
  - View and score Text Entry, Writing Prompts, multi-part item types.
  - Gain reporting information
- Start by locating the "Score" button for a completed session.

Class	Content Area	Test Name	Testing status	View	Delete	Score
Geo 1	Geometry	2017-2018 Pre-Test	In Progress	View		Score

41

Questar

NEW: Pre-Test Teacher Scoring

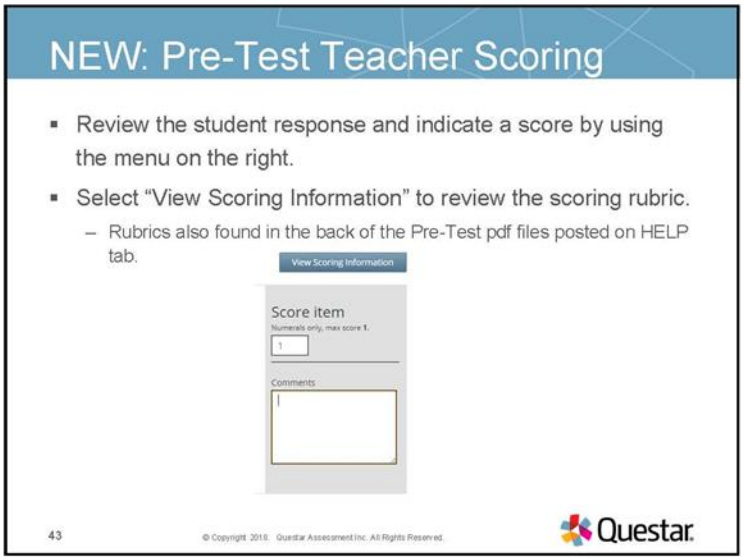
- Select "Score" next to a students session under the Hand Score column.

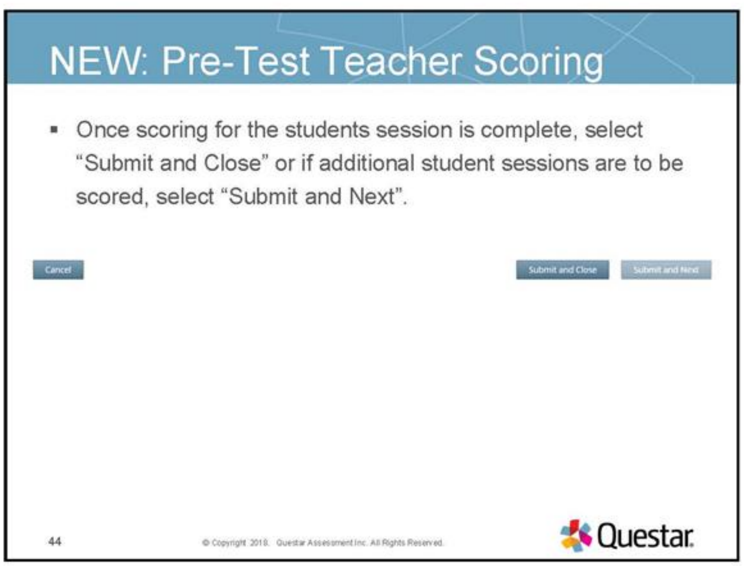
Testing status	Hand Score
Session 1: Finished	
Session 2: Finished	Score

- The entire class can be scored at one time, or each individual student.
- Student Pre-Test session must be in "Finished" status in order to score items.
  - Note: Not all sessions require Teacher scoring as many item types are machine scored.

42

Questar


NEW: Pre-Test Teacher Scoring
▪ Review the student response and indicate a score by using the menu on the right.
▪ Select "View Scoring Information" to review the scoring rubric.
– Rubrics also found in the back of the Pre-Test pdf files posted on HELP tab.
View Scoring Information
Score item
Numerals only, max score 1.
1
Comments
43
© Copyright 2018. Questar Assessment Inc. All Rights Reserved.
Questar


NEW: Pre-Test Teacher Scoring
▪ Once scoring for the students session is complete, select "Submit and Close" or if additional student sessions are to be scored, select "Submit and Next".
Cancel
Submit and Close
Submit and Next
44
© Copyright 2018. Questar Assessment Inc. All Rights Reserved.
Questar

Reports

- View and download reports by district, school, content area and report type.
- Report options will be based on the user's role (i.e. DTC, STC, or Teacher).

Questar

Help

- Access commonly asked questions
- View/Download Manuals, Quick Reference Guides, Training Modules
- Download and install New Secure Browser
- Contact Customer Support (MOCustomerSupport@QuestarAI.com)

Questar


Q&A

Appendix I: Accommodation Codes

ACCOMMODATIONS

These accommodations for use on the EOC Assessment are available only to students with an IEP/504 plan. Please read the full description prior to usage.

- All accommodations need to be marked in Nextera Admin prior to the assessment.
- Some accommodations are only for use by English Learner (EL) students (EL students are those marked LEP-RCV or LEP-NRC in Core Data).

Tool	Description	Code
Abacus	**Students with this accommodation in their IEP/504 plan** may have access to an abacus. This accommodation must be chosen in Nextera under student accommodations prior to testing.	A391
Alternate Response Options	**Students with this accommodation in their IEP/504 plan** may respond to items using an alternate option, including but not limited to: Adapted Keyboards, StickyKeys, MouseKeys, FilterKeys, Adapted Mouse, Touch Screen, Head Wand, and Switches. *Please Note: While the use of alternate response options is not directly supported by Questar, the help desk will work with districts needing to use one. The option must be provided by the district.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A441
Braille	**Students with visual impairments with this accommodation in their IEP/504 plan** may access the assessment via a Braille version. Tactile overlays and graphics tools may be used to assist the student in accessing the content. *Please Note: Answers from students who access the assessment using the Braille format must be entered into the Nextera student platform prior to shipping the Braille assessment back. Follow the instructions found in the Braille kit.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A012
Closed Captioning CC	Hearing Impaired students with this accommodation in their IEP/504 plan may have Closed Captioning available for ELA listening items. This accommodation must be chosen in Nextera under student accommodations prior to testing.	A053
Large Print	**Students with visual impairments with this accommodation in their IEP/504 plan** may access the assessment via a Large Print version. *Please Note: Answers from students who access the assessment using the Large Print format must be entered into the Nextera student platform prior to shipping the Large Print assessment back. Follow the instructions found in the Large Print kit.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A021

ACCOMMODATIONS

These accommodations for use on the EOC Assessment are available only to students with an IEP/504 plan. Please read the full description prior to usage.

- All accommodations need to be marked in Nextera Admin prior to the assessment.
- Some accommodations are only for use by English Learner (EL) students (EL students are those marked LEP-RCV or LEP-NRC in Core Data).

Tool	Description	Code
Multiplication Table	**Students with this accommodation in their IEP/504 plan** may have access to a single digit multiplication table. This accommodation must be chosen in Nextera under student accommodations prior to testing.	A395
Paper Based Assessment 	**Students with this accommodation in their IEP/504 plan** may take the assessment using the paper/pencil format. *Please Note: Answers from students who access the assessment using the Paper/Pencil format must be entered into the Nextera student platform prior to shipping the Paper assessment back.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A102
Read Aloud (ELA Reading Passages)	***Please see the Read Aloud section after the universal tools/ accommodations list.***	
Sign Language 	Hearing Impaired students with this accommodation in their IEP/504 plan may have ELA listening items translated into American Sign Language (ASL), Signing Exact English (SEE), or any other form of sign language. *Please Note: The Nextera Student Platform provides videos of ASL for these items. If the student uses SEE or another form of sign language, or the preference is for a local translation into ASL, the signing of ELA listening items will require the download of a script. See the manual for more details.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A052
Specialized Calculator	**Students with this accommodation in their IEP/504 plan** may have access to a specialized calculator. The specialized calculator can include a talking calculator or Braille calculator, among others. The memory of the physical calculator must be cleared before and after testing by the test examiner. *Please Note: Use of a calculator is only for the Mathematics and Science assessments.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A396

ACCOMMODATIONS These accommodations for use on the EOC Assessment are available only to students with an IEP/504 plan. Please read the full description prior to usage. • All accommodations need to be marked in Nextera Admin prior to the assessment. • Some accommodations are only for use by English Learner (EL) students (EL students are those marked LEP-RCV or LEP-NRC in Core Data).		
Tool	**Description**	**Code**
Speech-To-Text – Assistive Technology	**Students with this accommodation in their IEP/504 plan** may use that technology in conjunction with the Nextera student platform. The software must be provided by the district. *Please Note: The use of assistive technology software should be familiar to the student and should be software the student uses in the everyday classroom. While the use of assistive technology software is not directly supported by Questar, the help desk will work with districts needing to use the software. The software must be provided by the district.* This accommodation must be chosen in Nextera under student accommodations prior to testing.	A352

Appendix J. Classical Reliability Coefficients and Standard Error of Measurements

Table J.1. Alpha Coefficients and SEMs—English I, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	162	25.82	8.69	0.00	0.86	9.35
Gender						
Female	74	27.87	8.11	0.47	0.83	8.88
Male	88	24.07	8.83	0.00	0.87	9.46
Ethnicity						
American Indian/Alaskan Native	1	--	--	--	--	--
Asian	6	--	--	--	--	--
Black (not Hispanic)	33	--	--	0.19	0.83	8.43
Hispanic	27	--	--	-0.58	--	--
Multi-racial	4	--	--	--	--	--
White (not Hispanic)	91	26.27	8.39	0.00	0.86	9.05
LEP						
No	131	26.63	8.44	0.00	0.86	9.11
Yes	31	--	--	-0.46	--	--
IEP						
No	154	26.23	8.64	0.00	0.86	9.32
Yes	8	--	--	--	--	--
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	46	--	--	0.00	0.82	7.75
Yes	116	24.16	8.75	-0.67	0.86	9.42
Title I						
No	10	--	--	0.00	--	--
Yes	153	26.50	8.42	1.38	0.86	9.09
Accommodations						
No	143	26.58	8.71	0.00	0.86	9.37
Yes	19	--	--	-1.08	--	--

Table J.2. Alpha Coefficients and SEMs—English II, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	2,509	23.18	9.90	0.00	0.89	10.52
Gender						
Female	1,191	25.07	9.52	0.38	0.88	10.15
Male	1,310	21.49	9.94	0.00	0.89	10.57

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Ethnicity						
American Indian/Alaskan Native	9	--	--	--	--	
Asian	52	26.35	9.87	0.13	0.90	10.40
Black (not Hispanic)	540	18.34	7.81	-0.87	0.80	8.71
Hispanic	271	22.24	9.14	-0.31	0.85	9.89
Multi-racial	88	20.98	9.47	-0.44	0.86	10.19
Pacific Islander	9	--	--	--	--	--
White (not Hispanic)	1,533	25.11	10.09	0.00	0.89	10.68
LEP						
No	2,323	23.58	9.99	0.00	0.89	10.60
Yes	186	18.30	7.16	-0.74	0.74	8.32
IEP						
No	2,214	24.24	9.77	0.00	0.88	10.39
Yes	295	15.36	6.94	-1.28	0.77	7.93
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	1,165	26.28	10.29	0.00	0.90	10.87
Yes	1,345	20.52	8.72	-0.66	0.85	9.47
Title I						
No	2,004	23.55	10.12	0.00	0.89	10.72
Yes	507	21.75	8.88	-0.20	0.85	9.62
Accommodations						
No	2,244	23.92	9.89	0.00	0.89	10.51
Yes	265	16.96	7.56	-0.92	0.81	8.41

Table J.3. Alpha Coefficients and SEMs—Algebra I, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	5,271	21.26	10.70	0.00	0.92	11.17
Gender						
Female	2,495	21.76	10.56	0.08	0.91	11.04
Male	2,742	20.88	10.82	0.00	0.92	11.29
Ethnicity						
American Indian/Alaskan Native	20	--	--	-0.03	--	--
Asian	88	30.48	13.04	0.55	0.95	13.36
Black (not Hispanic)	1,017	15.17	7.49	-1.09	0.85	8.13
Hispanic	445	18.16	9.05	-0.57	0.88	9.63
Multi-racial	177	19.62	10.10	-0.37	0.91	10.59
Pacific Islander	11	--	--	-0.78	--	--
White (not Hispanic)	3,477	23.36	10.79	0.00	0.92	11.26

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
LEP						
No	5,037	21.55	10.74	0.00	0.92	11.21
Yes	234	15.12	7.40	-0.87	0.84	8.08
IEP						
No	4,678	22.20	10.68	0.00	0.92	11.15
Yes	593	13.91	7.55	-1.10	0.86	8.14
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	2,805	25.26	10.89	0.00	0.92	11.36
Yes	2,466	16.72	8.43	-1.01	0.87	9.02
Title I						
No	4,535	22.13	10.86	0.00	0.92	11.33
Yes	736	15.88	7.70	-0.81	0.84	8.40
Accommodations						
No	4,808	21.83	10.75	0.00	0.92	11.22
Yes	463	15.31	7.99	-0.82	0.87	8.59

Table J.4. Alpha Coefficients and SEMs—Algebra II, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	545	23.04	10.60	0.00	0.92	11.06
Gender						
Female	272	22.61	9.80	-0.10	0.90	10.32
Male	271	23.61	11.27	0.00	0.93	11.70
Ethnicity						
Asian	11	--	--	0.38	0.93	11.71
Black (not Hispanic)	44	--	--	-0.89	0.89	9.60
Hispanic	50	19.10	9.11	-0.56	--	--
Multi-racial	21	--	--	-0.14	0.94	12.03
White (not Hispanic)	417	24.20	10.42	0.00	0.92	10.89
LEP						
No	532	23.18	10.59	0.00	0.92	11.06
Yes	13	--	--	-0.63	--	--
IEP						
No	536	23.09	10.60	0.00	0.92	11.06
Yes	9	--	--	--	--	--
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
FRL						
No	376	24.67	10.36	0.00	0.92	10.82
Yes	169	19.41	10.25	-0.51	0.91	10.75
Title I						
No	479	24.17	10.45	0.00	0.92	10.92
Yes	66	14.82	7.64	-1.22	--	--
Accommodations						
No	539	23.04	10.60	0.00	0.92	11.06
Yes	6	--	--	--	--	--

Table J.5. Alpha Coefficients and SEMs—Geometry, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	137	31.80	9.22	0.00	0.91	9.68
Gender						
Female	71	30.46	9.57	-0.29	0.91	10.03
Male	66	33.27	8.64	0.00	0.89	9.14
Ethnicity						
Asian	5	--	--	--	--	--
Black (not Hispanic)	8	--	--	--	--	--
Hispanic	10	--	--	-1.11	--	--
Multi-racial	4	--	--	--	--	--
White (not Hispanic)	110	32.43	8.93	0.00	0.90	9.40
LEP						
No	137	31.80	9.22	0.00	0.91	9.68
Yes	--	--	--	--	--	--
IEP						
No	135	31.81	9.16	0.00	0.90	9.63
Yes	2	--	--	--	--	--
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	89	34.72	7.17	0.00	0.84	7.81
Yes	48	--	--	-0.82	0.92	10.61
Title I						
No	112	34.68	7.02	0.00	0.83	7.70
Yes	25	--	--	-2.53	--	--
Accommodations						
No	136	31.76	9.23	0.00	0.91	9.70
Yes	1	--	--	--	--	--

Table J.6. Alpha Coefficients and SEMs—Biology, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	1,315	26.20	11.67	0.00	0.92	12.16
Gender						
Female	650	27.09	11.46	0.15	0.92	11.95
Male	663	25.34	11.80	0.00	0.92	12.29
Ethnicity						
American Indian/Alaskan Native	5	--	--	--	--	--
Asian	39	--	--	0.44	0.93	11.55
Black (not Hispanic)	231	18.42	8.66	-1.16	0.84	9.42
Hispanic	115	23.98	10.78	-0.41	0.90	11.33
Multi-racial	62	27.64	11.63	-0.07	0.91	12.18
Pacific Islander	3	--	--	--	--	--
White (not Hispanic)	851	28.43	11.47	0.00	0.92	11.93
LEP						
No	1,258	26.52	11.71	0.00	0.92	12.20
Yes	57	19.35	8.18	-0.88	0.84	8.91
IEP						
No	1,160	27.24	11.55	0.00	0.92	12.03
Yes	155	18.65	9.63	-0.89	0.87	10.32
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	692	30.78	11.38	0.00	0.93	11.81
Yes	623	21.20	9.79	-0.98	0.88	10.45
Title I						
No	1,122	27.36	11.78	0.00	0.92	12.25
Yes	195	19.57	8.37	-0.93	0.82	9.23
Accommodations						
No	1,181	26.99	11.68	0.00	0.92	12.17
Yes	134	19.21	8.94	-0.87	0.86	9.66

Table J.7. Alpha Coefficients and SEMs—Biology, Fall 2018–Core Form B

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	1,186	26.20	11.67	0.00	0.93	12.10
Gender						
Female	582	27.09	11.46	0.15	0.93	11.90
Male	604	25.34	11.80	0.00	0.93	12.22
Ethnicity						
American Indian/Alaskan Native	8	--	--	--	--	--
Asian	37	--	--	0.44	0.93	11.55

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Black (not Hispanic)	239	18.42	8.66	-1.16	0.87	9.30
Hispanic	101	23.98	10.78	-0.41	0.91	11.27
Multi-racial	50	27.64	11.63	-0.07	0.94	12.02
Pacific Islander	3	--	--	--	--	--
White (not Hispanic)	743	28.43	11.47	0.00	0.93	11.89
LEP						
No	1,131	26.52	11.71	0.00	0.93	12.14
Yes	55	19.35	8.18	-0.88	0.83	8.97
IEP						
No	1,037	27.24	11.55	0.00	0.93	11.97
Yes	149	18.65	9.63	-0.89	0.90	10.15
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	612	30.78	11.38	0.00	0.93	11.80
Yes	574	21.20	9.79	-0.98	0.90	10.32
Title I						
No	1,005	27.36	11.78	0.00	0.93	12.20
Yes	183	19.57	8.37	-0.93	0.87	8.99
Accommodations						
No	1,066	26.99	11.68	0.00	0.93	12.10
Yes	120	19.21	8.94	-0.87	0.88	9.56

Table J.8. Alpha Coefficients and SEMs—Physical Science, Fall 2018–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	39	--	--	0.00	0.76	7.75
Gender						
Female	22	--	--	-0.27	0.69	7.40
Male	17	--	--	0.00	--	--
Ethnicity						
Black (not Hispanic)	1	--	--	--	--	--
Hispanic	12	--	--	-0.94	--	--
White (not Hispanic)	26	--	--	0.00	0.78	7.96
LEP						
No	35	--	--	0.00	0.75	7.65
Yes	4	--	--	--	--	--
IEP						
No	38	--	--	0.00	0.75	7.70
Yes	1	--	--	--	--	--

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
Title I						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
Accommodations						
No	39	--	--	0.00	0.76	7.75
Yes	--	--	--	--	--	--

Table J.9. Alpha Coefficients and SEMs—English I, Spring 2019–Core Form C

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	5,767	28.68	8.02	0.00	0.85	8.69
Gender						
Female	2,845	30.13	7.51	0.38	0.84	8.20
Male	2,916	27.25	8.24	0.00	0.85	8.92
Ethnicity						
American Indian/Alaskan Native	19	--	--	-0.45	--	--
Asian	68	30.02	8.89	0.04	0.86	9.57
Black (not Hispanic)	693	23.96	8.68	-0.66	0.85	9.41
Hispanic	407	26.61	8.34	-0.37	0.85	9.08
Multi-racial	145	28.14	7.43	-0.20	0.82	8.21
Pacific Islander	20	--	--	-0.27	--	--
White (not Hispanic)	4,409	29.66	7.55	0.00	0.84	8.22
LEP						
No	5,554	28.90	7.92	0.00	0.85	8.60
Yes	213	22.74	8.24	-0.75	0.82	9.07
IEP						
No	5,215	29.53	7.60	0.00	0.84	8.29
Yes	552	20.23	7.14	-1.30	0.78	8.10
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	2,650	31.07	7.06	0.00	0.83	7.75
Yes	3,117	26.68	8.22	-0.53	0.85	8.93
Title I						
No	4,787	29.46	7.59	0.00	0.84	8.27
Yes	980	25.01	8.93	-0.50	0.86	9.63
Accommodations						
No	4,867	29.31	7.77	0.00	0.84	8.45
Yes	900	23.90	8.28	-0.65	0.86	8.94

Table J.10. Alpha Coefficients and SEMs—English I, Spring 2019–Core Form D

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	5,098	28.68	8.02	0.00	0.83	8.78
Gender						
Female	2,584	30.13	7.51	0.38	0.82	8.29
Male	2,509	27.25	8.24	0.00	0.84	9.00
Ethnicity						
American Indian/Alaskan Native	19	--	--	-0.45	0.87	8.32
Asian	62	30.02	8.89	0.04	0.87	9.52

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Black (not Hispanic)	695	23.96	8.68	-0.66	0.84	9.48
Hispanic	287	26.61	8.34	-0.37	0.84	9.13
Multi-racial	119	28.14	7.43	-0.20	0.81	8.25
Pacific Islander	11	--	--	-0.27	--	--
White (not Hispanic)	3,900	29.66	7.55	0.00	0.82	8.35
LEP						
No	4,926	28.90	7.92	0.00	0.83	8.69
Yes	172	22.74	8.24	-0.75	0.82	9.09
IEP						
No	4,817	29.53	7.60	0.00	0.83	8.36
Yes	281	20.23	7.14	-1.30	0.78	8.07
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	2,341	31.07	7.06	0.00	0.80	7.89
Yes	2,757	26.68	8.22	-0.53	0.83	9.01
Title I						
No	4,184	29.46	7.59	0.00	0.82	8.37
Yes	915	25.01	8.93	-0.50	0.85	9.69
Accommodations						
No	4,873	29.31	7.77	0.00	0.83	8.52
Yes	225	23.90	8.28	-0.65	0.81	9.18

Table J.11. Alpha Coefficients and SEMs— English II, Spring 2019–Core Form C

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	32,562	28.53	8.32	0.00	0.87	8.93
Gender						
Female	16,185	29.83	7.87	0.32	0.86	8.48
Male	16,316	27.27	8.53	0.00	0.87	9.15
Ethnicity						
American Indian/Alaskan Native	126	28.07	7.66	-0.19	0.85	8.30
Asian	762	31.86	8.33	0.28	0.89	8.85
Black (not Hispanic)	4,762	23.86	8.07	-0.70	0.84	8.80
Hispanic	2,082	26.70	8.16	-0.34	0.86	8.81
Multi-racial	1,013	28.75	8.16	-0.09	0.86	8.81
Pacific Islander	83	25.04	7.37	-0.60	0.83	8.10
White (not Hispanic)	23,672	29.49	8.03	0.00	0.86	8.65
LEP						
No	31,555	28.74	8.26	0.00	0.87	8.88
Yes	1,007	21.71	7.21	-0.97	0.80	8.05

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
IEP						
No	29,025	29.55	7.83	0.00	0.86	8.45
Yes	3,537	20.07	7.39	-1.28	0.81	8.22
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	18,342	30.73	7.63	0.00	0.85	8.25
Yes	14,221	25.62	8.29	-0.62	0.86	8.97
Title I						
No	29,550	28.96	8.20	0.00	0.87	8.82
Yes	3,014	23.94	8.15	-0.62	0.85	8.84
Accommodations						
No	26,377	29.38	7.95	0.00	0.86	8.57
Yes	6,185	23.38	8.63	-0.69	0.87	9.26

Table J.12. Alpha Coefficients and SEMs— English II, Spring 2019–Core Form D

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	27,311	28.53	8.32	0.00	0.84	9.05
Gender						
Female	13,586	29.83	7.87	0.32	0.83	8.63
Male	13,680	27.27	8.53	0.00	0.85	9.26
Ethnicity						
American Indian/Alaskan Native	88	28.07	7.66	-0.19	0.80	8.56
Asian	612	31.86	8.33	0.28	0.86	8.96
Black (not Hispanic)	3,670	23.86	8.07	-0.70	0.82	8.91
Hispanic	1,683	26.70	8.16	-0.34	0.83	8.96
Multi-racial	872	28.75	8.16	-0.09	0.84	8.90
Pacific Islander	60	25.04	7.37	-0.60	0.80	8.23
White (not Hispanic)	20,281	29.49	8.03	0.00	0.84	8.78
LEP						
No	26,556	28.74	8.26	0.00	0.84	9.00
Yes	755	21.71	7.21	-0.97	0.75	8.31
IEP						
No	25,603	29.55	7.83	0.00	0.84	8.55
Yes	1,708	20.07	7.39	-1.28	0.80	8.23
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	15,984	30.73	7.63	0.00	0.83	8.40

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Yes	11,327	25.62	8.29	-0.62	0.84	9.07
Title I						
No	25,102	28.96	8.20	0.00	0.84	8.95
Yes	2,209	23.94	8.15	-0.62	0.83	8.96
Accommodations						
No	26,202	29.38	7.95	0.00	0.84	8.67
Yes	1,109	23.38	8.63	-0.69	0.82	9.52

Table J.13. Alpha Coefficients and SEMs—Algebra I, Spring 2019–Core Form C

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	32,031	22.31	9.68	0.00	0.89	10.24
Gender						
Female	15,936	22.57	9.46	0.05	0.89	10.04
Male	16,005	22.09	9.88	0.00	0.90	10.42
Ethnicity						
American Indian/Alaskan Native	140	19.97	9.45	-0.35	0.89	10.03
Asian	711	29.69	10.62	0.61	0.92	11.10
Black (not Hispanic)	4,714	17.18	8.30	-0.73	0.86	8.94
Hispanic	2,136	20.59	9.08	-0.29	0.88	9.70
Multi-racial	1,048	22.70	9.62	-0.06	0.89	10.18
Pacific Islander	88	19.82	8.45	-0.41	0.85	9.18
White (not Hispanic)	23,100	23.24	9.54	0.00	0.89	10.12
LEP						
No	30,830	22.46	9.68	0.00	0.89	10.24
Yes	1,201	18.45	8.84	-0.45	0.87	9.47
IEP						
No	28,121	23.30	9.46	0.00	0.89	10.05
Yes	3,910	14.29	7.48	-1.20	0.85	8.13
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	17,431	25.07	9.60	0.00	0.89	10.18
Yes	14,600	18.90	8.62	-0.72	0.87	9.25
Title I						
No	28,500	22.84	9.65	0.00	0.89	10.22
Yes	3,531	17.72	8.64	-0.59	0.87	9.25
Accommodations						
No	25,801	23.23	9.51	0.00	0.89	10.10
Yes	6,230	16.47	8.63	-0.78	0.87	9.23

Table J.14. Alpha Coefficients and SEMs—Algebra I, Spring 2019–Core Form D

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	26,451	22.31	9.68	0.00	0.88	10.31
Gender						
Female	13,130	22.57	9.46	0.05	0.87	10.12
Male	13,239	22.09	9.88	0.00	0.89	10.49
Ethnicity						
American Indian/Alaskan Native	87	19.97	9.45	-0.35	0.86	10.17
Asian	629	29.69	10.62	0.61	0.91	11.12
Black (not Hispanic)	3,434	17.18	8.30	-0.73	0.85	8.99
Hispanic	1,692	20.59	9.08	-0.29	0.87	9.74
Multi-racial	980	22.70	9.62	-0.06	0.88	10.27
Pacific Islander	78	19.82	8.45	-0.41	0.87	9.04
White (not Hispanic)	19,466	23.24	9.54	0.00	0.88	10.20
LEP						
No	25,513	22.46	9.68	0.00	0.88	10.32
Yes	938	18.45	8.84	-0.45	0.87	9.46
IEP						
No	24,832	23.30	9.46	0.00	0.88	10.10
Yes	1,619	14.29	7.48	-1.20	0.85	8.09
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	15,026	25.07	9.60	0.00	0.88	10.24
Yes	11,425	18.90	8.62	-0.72	0.85	9.34
Title I						
No	23,844	22.84	9.65	0.00	0.88	10.29
Yes	2,607	17.72	8.64	-0.59	0.86	9.32
Accommodations						
No	25,622	23.23	9.51	0.00	0.88	10.14
Yes	829	16.47	8.63	-0.78	0.83	9.49

Table J.15. Alpha Coefficients and SEMs—Algebra II, Spring 2019–Core Form C

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	8,062	24.93	9.30	0.00	0.88	9.93
Gender						
Female	4,267	24.58	9.04	-0.08	--	--
Male	3,777	25.34	9.57	0.00	0.88	10.17
Ethnicity						
American Indian/Alaskan Native	22	--	--	-0.03	--	--
Asian	352	31.20	9.63	0.62	--	--

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Black (not Hispanic)	613	19.97	8.43	-0.62	--	--
Hispanic	419	23.33	8.88	-0.21	--	--
Multi-racial	266	24.29	9.43	-0.10	--	--
Pacific Islander	11	--	--	-0.25	--	--
White (not Hispanic)	6,360	25.21	9.13	0.00	0.87	9.78
LEP						
No	7,930	24.99	9.28	0.00	0.88	9.91
Yes	132	21.19	9.59	-0.40	--	--
IEP						
No	7,939	25.00	9.27	0.00	0.88	9.91
Yes	123	19.89	9.73	-0.53		
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	5,898	26.20	9.21	0.00	0.88	9.83
Yes	2,164	21.45	8.63	-0.55	--	--
Title I						
No	7,665	25.16	9.28	0.00	0.88	9.91
Yes	397	20.02	8.21	-0.63	--	--
Accommodations						
No	7,227	25.02	9.31	0.00	0.88	9.94
Yes	835	23.49	8.92	-0.17	--	--

Table J.16. Alpha Coefficients and SEMs—Algebra II, Spring 2019–Core Form D

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	7,501	24.93	9.30	0.00	0.88	9.90
Gender						
Female	4,015	24.58	9.04	-0.08	0.88	9.67
Male	3,469	25.34	9.57	0.00	0.89	10.15
Ethnicity						
American Indian/Alaskan Native	16	--	--	-0.03	--	--
Asian	276	31.20	9.63	0.62	0.88	10.27
Black (not Hispanic)	508	19.97	8.43	-0.62	0.87	9.06
Hispanic	395	23.33	8.88	-0.21	0.87	9.49
Multi-racial	233	24.29	9.43	-0.10	0.88	10.03
Pacific Islander	10	--	--	-0.25	0.95	13.35
White (not Hispanic)	6,046	25.21	9.13	0.00	0.88	9.74
LEP						
No	7,381	24.99	9.28	0.00	0.88	9.88
Yes	120	21.19	9.59	-0.40	0.89	10.19

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
IEP						
No	7,421	25.00	9.27	0.00	0.88	9.88
Yes	80	19.89	9.73	-0.53	0.91	10.22
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	5,511	26.20	9.21	0.00	0.88	9.82
Yes	1,990	21.45	8.63	-0.55	0.87	9.28
Title I						
No	7,198	25.16	9.28	0.00	0.88	9.89
Yes	303	20.02	8.21	-0.63	0.87	8.82
Accommodations						
No	7,464	25.02	9.31	0.00	0.88	9.92
Yes	37	--	--	-0.17	0.87	9.56

Table J.17. Alpha Coefficients and SEMs—Geometry, Spring 2019–Core Form C

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	1,837	22.03	8.97	0.00	0.87	9.60
Gender						
Female	949	21.73	8.63	-0.07	0.86	9.28
Male	879	22.31	9.30	0.00	0.88	9.91
Ethnicity						
American Indian/Alaskan Native	6	--	--	--	--	--
Asian	36	--	--	1.46	0.89	9.68
Black (not Hispanic)	64	20.06	7.73	-0.24	0.80	8.67
Hispanic	87	20.06	7.86	-0.23	0.86	8.48
Multi-racial	44	--	--	-0.01	0.88	8.38
Pacific Islander	2	--	--	--	--	--
White (not Hispanic)	1,589	21.89	8.84	0.00	0.87	9.49
LEP						
No	1,791	22.10	8.99	0.00	0.87	9.62
Yes	46	--	--	-0.41	0.81	8.24
IEP						
No	1,741	22.28	8.96	0.00	0.87	9.59
Yes	96	16.48	7.26	-0.80	0.81	8.09
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	1,115	23.52	9.36	0.00	0.88	9.96

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Yes	722	19.53	7.64	-0.52	0.83	8.37
Title I						
No	1,723	22.14	9.03	0.00	0.87	9.66
Yes	114	20.19	7.74	-0.25	0.85	8.42
Accommodations						
No	1,735	22.25	8.97	0.00	0.87	9.60
Yes	102	17.31	7.53	-0.66	0.84	8.22

Table J.18. Alpha Coefficients and SEMs—Geometry, Spring 2019–Core Form D

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	1,718	22.03	8.97	0.00	0.86	9.65
Gender						
Female	888	21.73	8.63	-0.07	0.85	9.37
Male	825	22.31	9.30	0.00	0.88	9.94
Ethnicity						
American Indian/Alaskan Native	1	--	--	--	--	--
Asian	36	--	--	1.46	0.90	9.64
Black (not Hispanic)	58	20.06	7.73	-0.24	0.83	8.50
Hispanic	74	20.06	7.86	-0.23	0.82	8.70
Multi-racial	43	--	--	-0.01	0.75	9.05
Pacific Islander	4	--	--	--	--	--
White (not Hispanic)	1,497	21.89	8.84	0.00	0.86	9.55
LEP						
No	1,684	22.10	8.99	0.00	0.86	9.67
Yes	34	--	--	-0.41	0.81	8.25
IEP						
No	1,678	22.28	8.96	0.00	0.86	9.65
Yes	40	--	--	-0.80	0.84	7.91
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	1,117	23.52	9.36	0.00	0.87	10.02
Yes	601	19.53	7.64	-0.52	0.81	8.50
Title I						
No	1,631	22.14	9.03	0.00	0.87	9.70
Yes	87	20.19	7.74	-0.25	0.77	8.81
Accommodations						
No	1,678	22.25	8.97	0.00	0.86	9.66
Yes	40	--	--	-0.66	--	--

Table J.19. Alpha Coefficients and SEMs—Biology, Spring 2019–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	33,873	30.30	10.16	0.00	0.90	10.69
Gender						
Female	16,596	30.49	9.79	0.04	0.90	10.34
Male	17,255	30.11	10.51	0.00	0.91	11.02
Ethnicity						
American Indian/Alaskan Native	143	29.68	9.76	-0.21	0.90	10.30
Asian	720	35.05	10.07	0.33	0.92	10.49
Black (not Hispanic)	5,115	23.62	9.76	-0.83	0.88	10.41
Hispanic	2,081	27.85	10.03	-0.39	0.89	10.63
Multi-racial	1,106	30.43	10.00	-0.13	0.90	10.53
Pacific Islander	100	25.59	9.38	-0.65	0.88	9.97
White (not Hispanic)	24,552	31.74	9.64	0.00	0.90	10.18
LEP						
No	32,740	30.56	10.08	0.00	0.90	10.60
Yes	1,133	22.31	9.49	-0.87	0.87	10.19
IEP						
No	28,943	31.36	9.70	0.00	0.89	10.26
Yes	4,930	21.34	9.48	-1.06	0.87	10.17
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	18,848	33.17	9.25	0.00	0.90	9.78
Yes	15,025	26.53	10.07	-0.66	0.89	10.67
Title I						
No	30,555	30.95	9.96	0.00	0.90	10.48
Yes	3,321	24.19	10.01	-0.68	0.89	10.63
Accommodations						
No	25,912	31.25	9.83	0.00	0.90	10.38
Yes	7,961	24.73	10.30	-0.63	0.90	10.87

Table J.20. Alpha Coefficients and SEMs—Biology, Spring 2019–Core Form B

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	26,533	30.30	10.16	0.00	0.91	10.66
Gender						
Female	13,301	30.49	9.79	0.04	0.90	10.31
Male	13,216	30.11	10.51	0.00	0.91	10.99
Ethnicity						
American Indian/Alaskan Native	84	29.68	9.76	-0.21	0.89	10.36
Asian	606	35.05	10.07	0.33	0.92	10.52

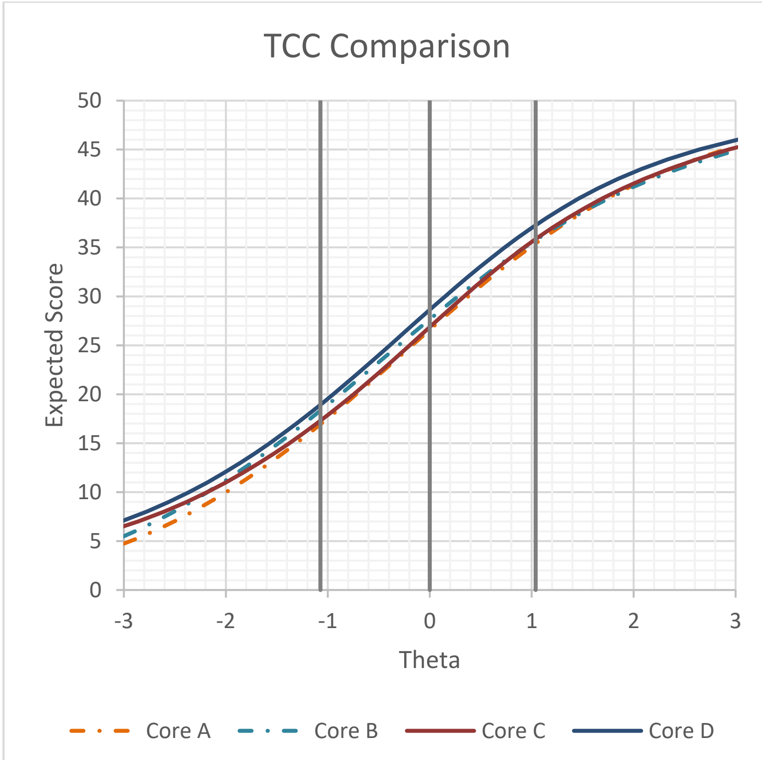
Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
Black (not Hispanic)	3,746	23.62	9.76	-0.83	0.89	10.34
Hispanic	1,698	27.85	10.03	-0.39	0.90	10.57
Multi-racial	846	30.43	10.00	-0.13	0.90	10.51
Pacific Islander	60	25.59	9.38	-0.65	0.87	10.04
White (not Hispanic)	19,448	31.74	9.64	0.00	0.90	10.17
LEP						
No	25,739	30.56	10.08	0.00	0.91	10.58
Yes	794	22.31	9.49	-0.87	0.89	10.07
IEP						
No	25,045	31.36	9.70	0.00	0.90	10.21
Yes	1,488	21.34	9.48	-1.06	0.90	10.00
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	15,462	33.17	9.25	0.00	0.89	9.79
Yes	11,071	26.53	10.07	-0.66	0.90	10.61
Title I						
No	24,017	30.95	9.96	0.00	0.90	10.47
Yes	2,517	24.19	10.01	-0.68	0.90	10.56
Accommodations						
No	25,663	31.25	9.83	0.00	0.91	10.32
Yes	870	24.73	10.30	-0.63	0.90	10.89

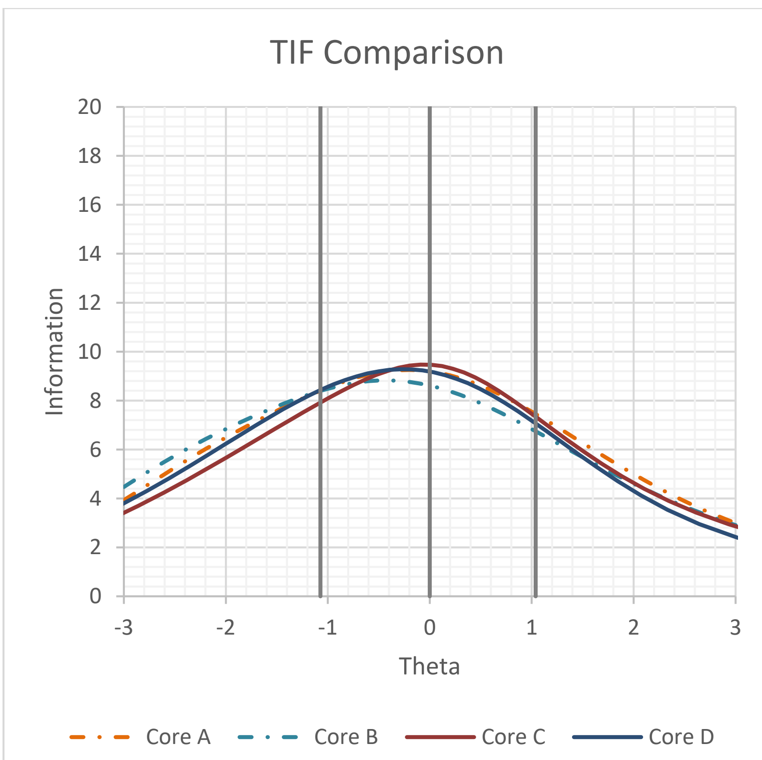
Table J.21. Alpha Coefficients and SEMs—Physical Science, Spring 2019–Core Form A

Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
All Students	2,286	27.17	8.79	0.00	0.86	9.46
Gender						
Female	1,105	26.72	8.40	-0.11	0.85	9.12
Male	1,181	27.60	9.12	0.00	0.88	9.74
Ethnicity						
American Indian/Alaskan Native	8	--	--	--	--	--
Asian	15	--	--	0.06	0.91	10.72
Black (not Hispanic)	120	21.43	7.93	-0.78	0.83	8.73
Hispanic	54	25.13	8.17	-0.31	0.84	8.91
Multi-racial	53	24.21	9.01	-0.38	0.86	9.69
Pacific Islander	2	--	--	--	--	--
White (not Hispanic)	2,032	27.65	8.70	0.00	0.86	9.38
LEP						
No	2,267	27.21	8.78	0.00	0.86	9.45
Yes	19	--	--	-0.53	0.84	9.06

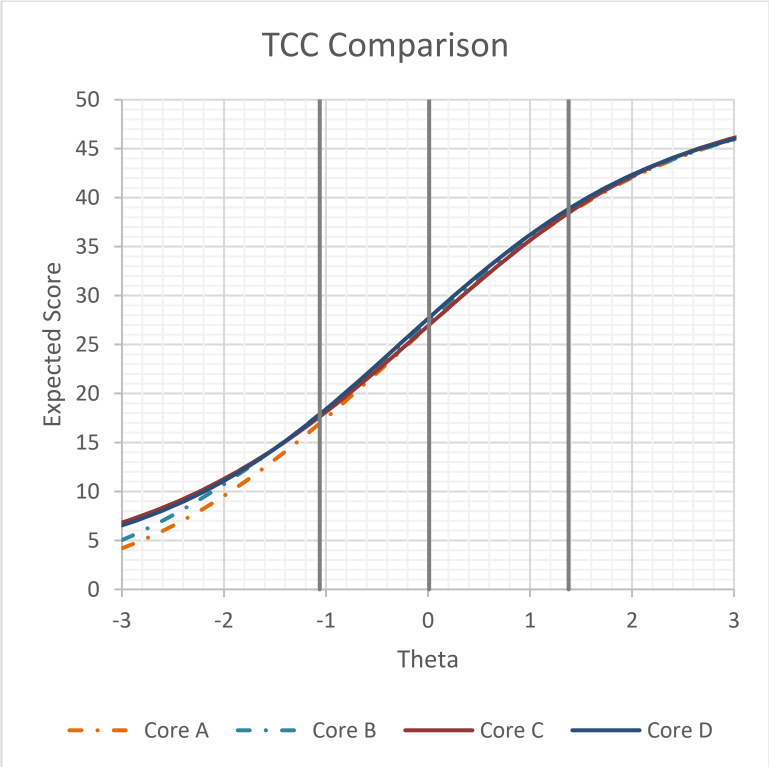
Group	*n*-Count	Mean Raw Score	SD Raw Score	Effect Size	Reliability	SEM
IEP						
No	2,113	27.88	8.58	0.00	0.86	9.27
Yes	173	18.52	6.36	-1.47	0.74	7.37
Migrant						
No	--	--	--	--	--	--
Yes	--	--	--	--	--	--
FRL						
No	1,292	29.21	8.54	0.00	0.86	9.22
Yes	994	24.53	8.40	-0.56	0.85	9.12
Title I						
No	2,168	27.28	8.83	0.00	0.87	9.50
Yes	118	25.21	7.63	-0.27	0.81	8.49
Accommodations						
No	2,123	27.78	8.61	0.00	0.86	9.30
Yes	163	19.33	7.16	-1.18	0.80	8.02

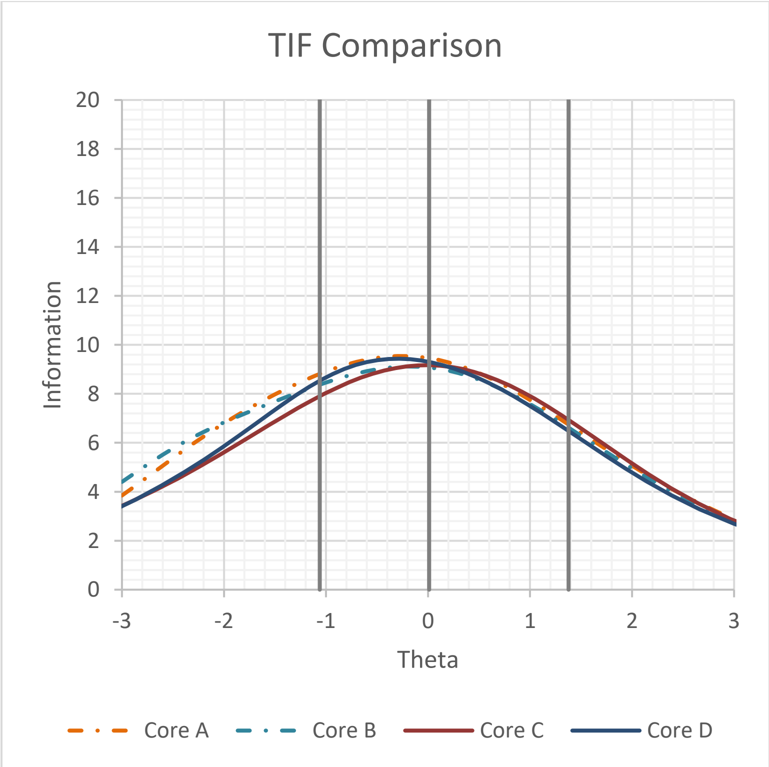
Appendix K: TCC and Conditional Standard Error
English I



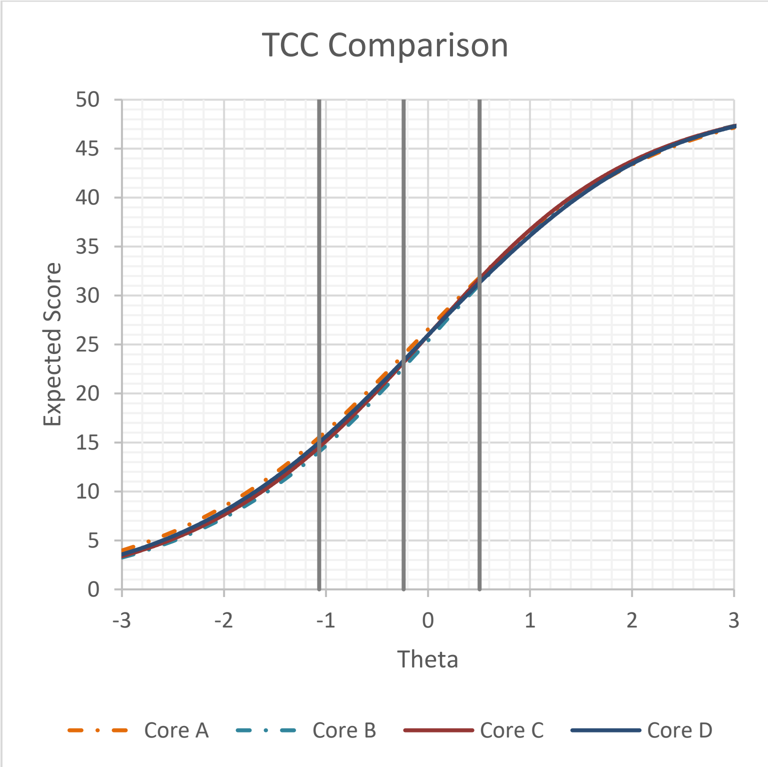


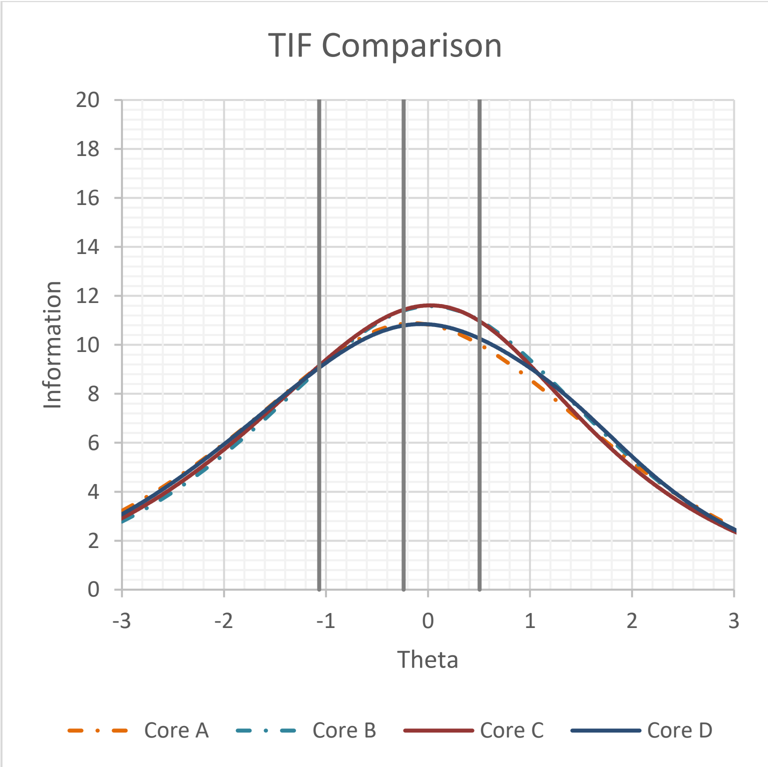
English II


TCC Comparison
Expected Score
Theta
Core A
Core B
Core C
Core D

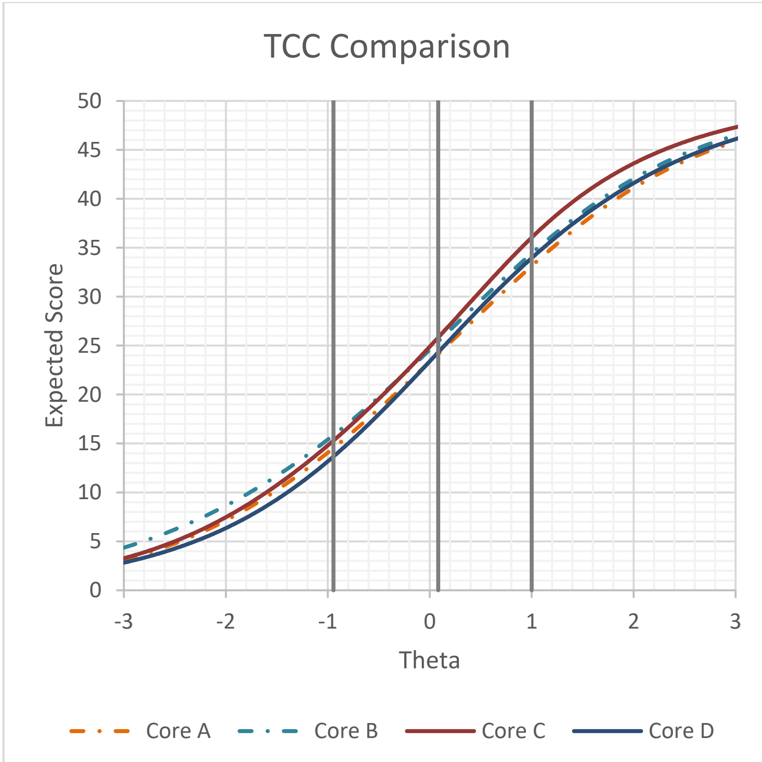

TIF Comparison
Information
Theta
Core A
Core B
Core C
Core D

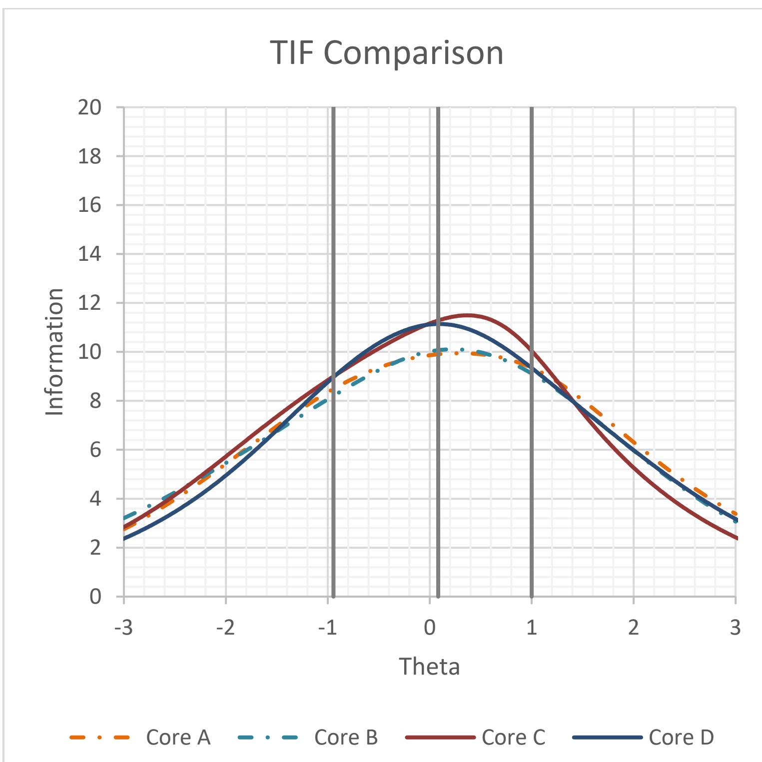
Algebra I


TCC Comparison
Expected Score
50
45
40
35
30
25
20
15
10
5
0
-3
-2
-1
0
1
2
3
Theta
Core A
Core B
Core C
Core D

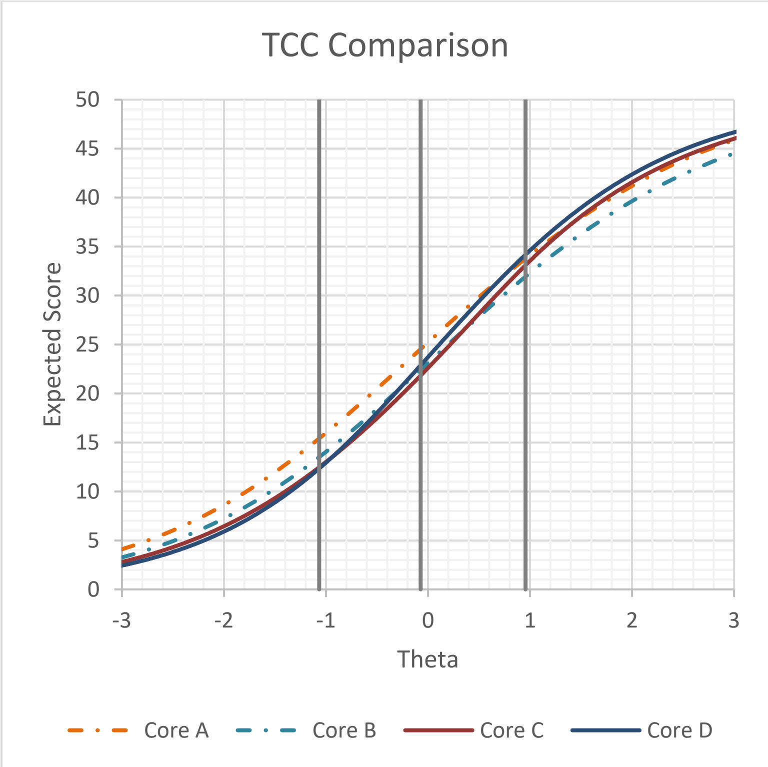

TIF Comparison
Information
20
18
16
14
12
10
8
6
4
2
0
-3
-2
-1
0
1
2
3
Theta
Core A
Core B
Core C
Core D

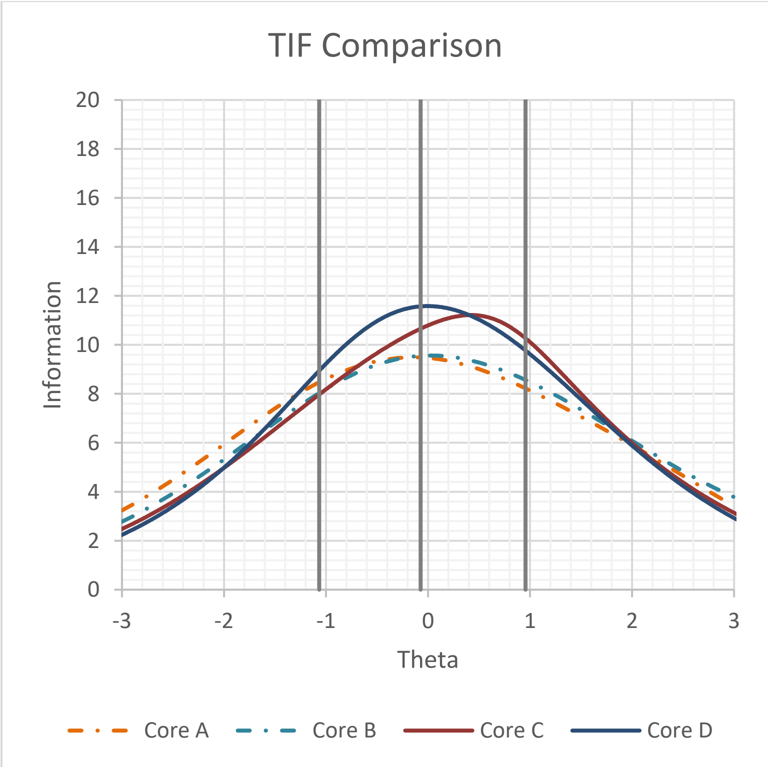
Algebra II


TCC Comparison
Expected Score
50
45
40
35
30
25
20
15
10
5
0
-3
-2
-1
0
1
2
3
Theta
Core A
Core B
Core C
Core D

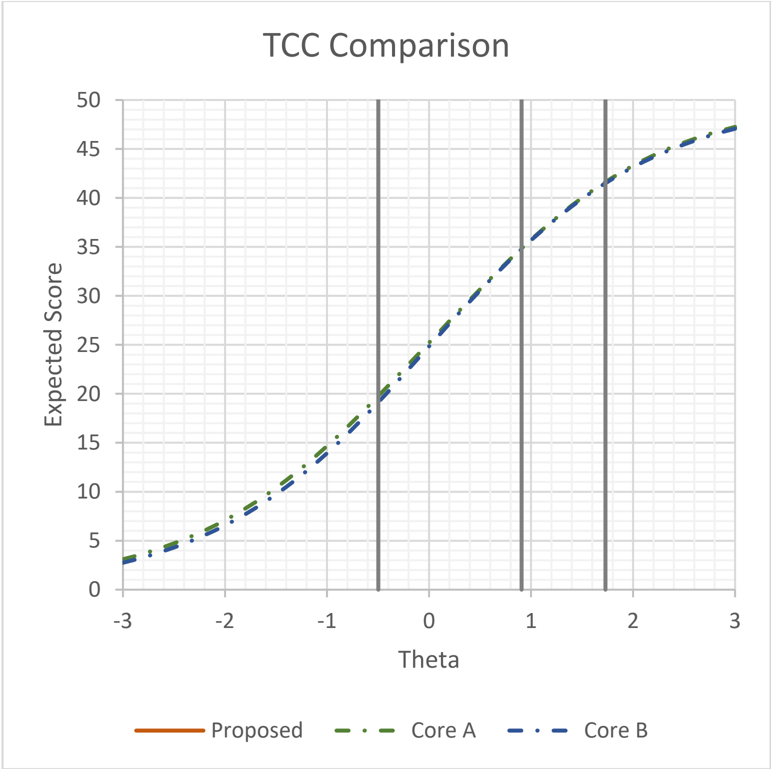

TIF Comparison
Information
20
18
16
14
12
10
8
6
4
2
0
-3
-2
-1
0
1
2
3
Theta
Core A
Core B
Core C
Core D

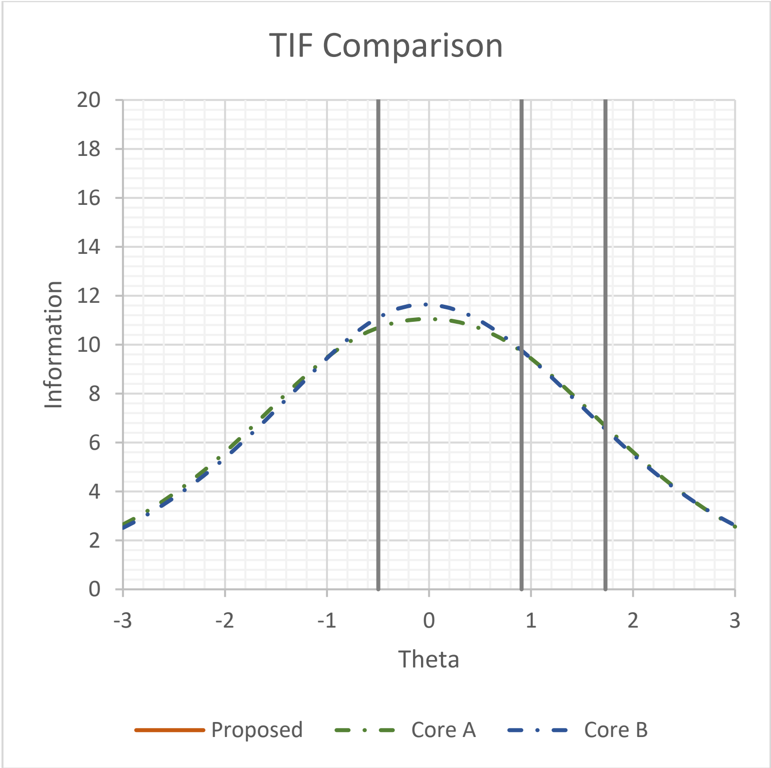
Geometry


TCC Comparison
Expected Score
50
45
40
35
30
25
20
15
10
5
0
-3
-2
-1
0
1
2
3
Theta
Core A
Core B
Core C
Core D

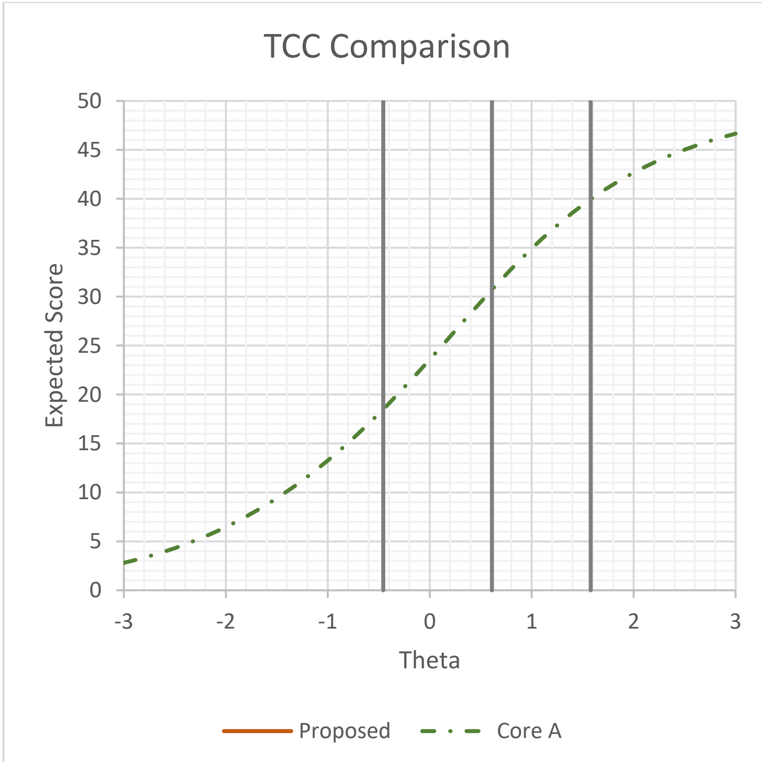

TIF Comparison
Information
20
18
16
14
12
10
8
6
4
2
0
-3
-2
-1
0
1
2
3
Theta
Core A
Core B
Core C
Core D

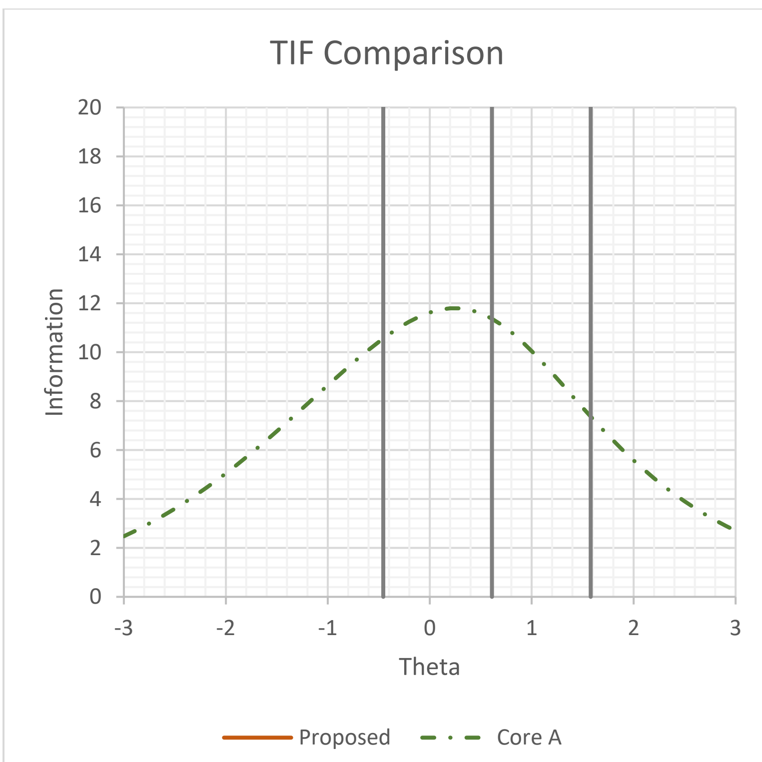
Biology


TCC Comparison
Expected Score
50
45
40
35
30
25
20
15
10
5
0
-3
-2
-1
0
1
2
3
Theta
Proposed
Core A
Core B


TIF Comparison
Information
20
18
16
14
12
10
8
6
4
2
0
-3
-2
-1
0
1
2
3
Theta
Proposed
Core A
Core B

Physical Science





Appendix L: IRT Item Statistics

Table L.1. IRT Item Statistics–English I–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0001782	-0.4087	0.0219	1.00	1.01	0.39	C & D
MO0001783	0.2106	0.0308	1.14	1.18	0.22	D
MO0001812	-1.8711	0.0294	0.88	0.80	0.44	C & D
MO0001818	1.0837	0.0310	0.95	0.88	0.42	C
MO0007208	-1.2932	0.0371	0.88	0.74	0.49	D
MO0007308	0.3914	0.0308	1.07	1.12	0.29	D
MO0007794	0.3258	0.0292	1.21	1.29	0.18	C
MO0007858	0.0858	0.0292	1.02	1.03	0.38	C
MO0007902	-0.1268	0.0295	0.91	0.89	0.49	C
MO0008149	0.1533	0.0292	0.99	0.99	0.41	C
MO0008157	-0.6175	0.0308	0.94	0.89	0.46	C
MO0008223	-0.4052	0.0301	1.06	1.07	0.34	C
MO0008224	0.0930	0.0308	0.92	0.91	0.46	D
MO0008256	0.1301	0.0308	1.07	1.09	0.30	D
MO0008285	0.6491	0.0312	0.95	0.97	0.41	D
MO0008334	0.1080	0.0292	0.99	0.99	0.41	C
MO0008336	0.9214	0.0304	0.92	0.95	0.44	C
MO0008337	0.0215	0.0293	1.03	1.04	0.37	C
MO0008349	-0.3872	0.0300	0.97	0.95	0.43	C
MO0008423	0.1272	0.0308	0.99	0.99	0.38	D
MO0008735	0.4789	0.0309	1.05	1.07	0.31	D
MO0008772	-0.0984	0.0311	0.98	0.98	0.39	D
MO0008780	-0.0275	0.0293	1.12	1.14	0.29	C
MO0008792	0.4094	0.0308	1.17	1.24	0.18	D
MO0008804	-1.2782	0.0370	0.91	0.82	0.45	D
MO0008908	-0.5602	0.0305	1.10	1.17	0.29	C
MO0018251	-0.4433	0.0320	1.02	1.02	0.35	D
MO0018258	0.1177	0.0308	1.10	1.14	0.26	D
MO0018266	-1.1110	0.0356	0.87	0.78	0.49	D
MO0018281	-0.8621	0.0340	0.90	0.86	0.47	D
MO0018534	-0.4290	0.0319	0.96	0.96	0.41	D
MO0044055	0.0311	0.0309	1.03	1.04	0.34	D
MO0044057	-1.1405	0.0359	0.94	0.88	0.42	D
MO0044058	-1.2208	0.0340	0.94	0.93	0.42	C
MO0044060	0.3487	0.0308	1.04	1.08	0.32	D
MO0044069	-1.1967	0.0338	0.83	0.69	0.55	C
MO0044075	-0.5157	0.0304	1.04	1.05	0.36	C
MO0044077	-1.2932	0.0371	0.98	0.95	0.37	D
MOE11614	-0.6922	0.0331	0.87	0.83	0.45	D
MOE116145	3.8640	0.0540	1.02	1.10	0.12	C & D
MOE116149	-0.4837	0.0303	1.02	1.02	0.41	C

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MOE116152	-0.7196	0.0227	1.02	1.01	0.36	C & D
MOE116161	0.0853	0.0308	0.95	0.96	0.41	D
MOE11617	0.1773	0.0212	0.94	0.94	0.44	C & D
MOE116223	2.2418	0.0401	1.03	1.07	0.23	C
MOE116298	1.3164	0.0322	0.98	1.02	0.34	C
MOE116299	-1.1538	0.0335	0.92	0.87	0.41	C
MOE116441	-0.1853	0.0296	0.99	0.98	0.40	C
MOE11675	-0.1257	0.0311	1.04	1.05	0.34	D
MOE11678	-0.1081	0.0311	1.11	1.14	0.28	D
MOE11680	0.1527	0.0308	1.06	1.07	0.32	D
MOE11688	-0.5299	0.0222	0.90	0.87	0.48	C & D
MOE11689	-0.1467	0.0214	1.07	1.10	0.31	C & D

Table L.2. IRT Item Statistics–English II–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0001825	0.3666	0.0090	1.07	1.12	0.33	C & D
MO0001831	-0.8232	0.0098	0.90	0.84	0.48	C & D
MO0001834	-0.7892	0.0097	0.91	0.86	0.47	C & D
MO0007288	0.6384	0.0133	1.24	1.32	0.13	D
MO0007627	0.1547	0.0132	0.94	0.93	0.45	D
MO0007733	-0.5712	0.0140	1.11	1.16	0.26	D
MO0007879	-0.5526	0.0139	0.95	0.91	0.43	D
MO0007941	-0.8934	0.0147	0.92	0.87	0.45	D
MO0008025	0.9220	0.0137	0.97	1.00	0.40	D
MO0008042	-1.2298	0.0157	1.09	1.27	0.23	D
MO0008076	-0.2741	0.0135	1.07	1.10	0.31	D
MO0008121	-0.2694	0.0125	0.97	0.94	0.44	C
MO0008177	0.0546	0.0133	0.89	0.87	0.50	D
MO0008207	0.9057	0.0136	1.10	1.23	0.25	D
MO0008211	0.6236	0.0124	1.11	1.15	0.30	C
MO0008236	0.1497	0.0123	0.95	0.94	0.47	C
MO0008352	0.0112	0.0123	0.88	0.86	0.53	C
MO0008485	0.7882	0.0135	0.99	1.04	0.38	D
MO0008495	-0.2795	0.0135	1.10	1.13	0.28	D
MO0008716	-0.1845	0.0124	0.95	0.96	0.46	C
MO0008742	-0.3543	0.0126	0.93	0.90	0.48	C
MO0008757	-1.0842	0.0139	0.88	0.79	0.50	C
MO0008765	-0.6349	0.0130	0.97	0.93	0.44	C
MO0008791	-0.3544	0.0126	0.96	0.94	0.45	C
MO0008813	1.3610	0.0135	1.17	1.44	0.18	C
MO0008827	0.7083	0.0125	1.21	1.29	0.20	C
MO0008859	-1.2136	0.0142	1.02	1.05	0.35	C
MO0008867	0.1604	0.0123	1.14	1.19	0.28	C

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0008877	0.4655	0.0123	0.97	1.00	0.43	C
MO0008956	0.3768	0.0123	1.00	1.03	0.41	C
MO0008971	1.2321	0.0132	1.16	1.37	0.20	C
MO0009087	-1.7144	0.0160	1.00	1.12	0.32	C
MO0009090	1.5667	0.0140	1.07	1.35	0.24	C
MO0044091	-0.7958	0.0144	0.89	0.83	0.48	D
MO0044092	-0.2920	0.0125	0.84	0.78	0.57	C
MO0044101	0.0038	0.0133	1.05	1.05	0.34	D
MOE2161	0.6166	0.0091	1.05	1.08	0.34	C & D
MOE21612	1.4201	0.0100	1.00	1.03	0.37	C & D
MOE216237	0.3115	0.0090	0.99	0.99	0.42	C & D
MOE216303	0.1280	0.0090	1.08	1.10	0.32	C & D
MOE2165	-0.4054	0.0093	1.03	1.03	0.36	C & D
MOE2166	0.5557	0.0091	1.10	1.18	0.29	C & D
MOE216709	-1.2245	0.0105	0.77	0.63	0.51	C & D
MOE21675	-1.0604	0.0102	1.08	1.12	0.36	C & D
MOE21678	-0.4022	0.0093	0.97	0.93	0.45	C & D
MOE216788	-0.4843	0.0094	0.86	0.81	0.53	C & D
MOE216791	0.9664	0.0094	0.92	0.86	0.47	C & D
MOE216793	0.3736	0.0132	0.97	0.95	0.42	D
MOE21683	-0.6241	0.0095	0.96	0.92	0.45	C & D

Table L.3. IRT Item Statistics–Algebra I–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0003373	0.4434	0.0132	0.99	1.05	0.42	C
MO0003682	-0.7195	0.0093	0.95	0.93	0.45	C & D
MO0003888	0.8403	0.0140	1.10	1.28	0.30	C
MO0007286	2.5331	0.0162	0.89	0.59	0.41	C & D
MO0007747	0.2539	0.0130	1.15	1.26	0.28	C
MO0007829	-2.3288	0.0123	1.03	1.24	0.25	C & D
MO0007928	0.2591	0.0095	1.15	1.26	0.27	C & D
MO0007929	0.2081	0.0095	1.13	1.19	0.29	C & D
MO0008117	0.3541	0.0131	1.07	1.11	0.36	C
MO0008160	-0.0979	0.0093	1.05	1.08	0.37	C & D
MO0008164	-0.3689	0.0137	0.84	0.80	0.56	D
MO0008175	-1.2169	0.0098	0.85	0.75	0.53	C & D
MO0008200	0.9584	0.0155	1.07	1.20	0.31	D
MO0008228	0.0774	0.0094	1.10	1.12	0.33	C & D
MO0008294	-0.1899	0.0137	0.89	0.86	0.52	D
MO0008313	-1.1056	0.0143	0.83	0.76	0.54	D
MO0008325	0.4253	0.0132	0.85	0.84	0.54	C
MO0008357	1.8246	0.0190	0.83	0.60	0.52	D
MO0008366	-0.3832	0.0137	1.09	1.11	0.32	D

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0008367	-1.2660	0.0098	0.89	0.85	0.48	C & D
MO0008370	-0.3180	0.0137	0.88	0.86	0.51	D
MO0008379	0.5485	0.0134	0.93	0.86	0.49	C
MO0008405	0.2429	0.0095	1.00	1.03	0.41	C & D
MO0008414	-0.2194	0.0093	1.05	1.07	0.37	C & D
MO0008437	-1.2307	0.0132	0.78	0.70	0.59	C
MO0008446	0.8582	0.0141	0.92	0.94	0.47	C
MO0008463	-0.1756	0.0126	1.21	1.28	0.25	C
MO0008476	-2.0208	0.0170	0.98	1.00	0.31	D
MO0008481	0.2153	0.0140	1.04	1.07	0.37	D
MO0008491	1.5543	0.0163	0.88	0.68	0.48	C
MO0008494	-0.4590	0.0126	1.09	1.12	0.34	C
MO0008519	1.8034	0.0175	0.98	1.13	0.34	C
MO0008715	-0.1054	0.0138	1.06	1.09	0.35	D
MO0008747	-0.1107	0.0127	0.85	0.81	0.56	C
MO0008749	-1.1414	0.0097	1.11	1.18	0.28	C & D
MO0008774	-0.1040	0.0138	1.07	1.09	0.35	D
MO0008794	0.1386	0.0129	0.90	0.84	0.52	C
MO0008811	-0.2530	0.0137	0.95	0.93	0.46	D
MO0008893	-0.8726	0.0140	1.01	1.01	0.38	D
MO0008907	-1.7320	0.0143	1.12	1.40	0.23	C
MO0008961	-0.1459	0.0093	1.04	1.05	0.39	C & D
MO0008965	-0.0394	0.0093	0.90	0.87	0.51	C & D
MO0008969	-0.0338	0.0094	1.05	1.04	0.38	C & D
MO0009099	0.5835	0.0146	0.80	0.74	0.59	D
MO0009146	1.1417	0.0148	1.02	1.17	0.35	C
MO0009204	-1.8025	0.0161	0.90	0.79	0.42	D
MO0009213	-0.0360	0.0138	1.02	1.03	0.39	D
MO0009217	-1.1695	0.0131	0.94	0.93	0.45	C
MO0009230	0.5570	0.0146	1.07	1.14	0.33	D
MO0009238	0.0123	0.0127	1.18	1.28	0.27	C
MO0009242	-1.5454	0.0103	0.91	0.83	0.45	C & D
MOA116159	0.1040	0.0139	1.00	1.01	0.41	D

Table L.4. IRT Item Statistics–Algebra II–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0003376	1.3994	0.0204	0.89	0.82	0.49	C & D
MO0003378	-0.7429	0.0270	0.85	0.79	0.52	D
MO0007276	-0.2271	0.0259	0.84	0.80	0.55	D
MO0007277	0.0860	0.0249	0.93	0.89	0.48	C
MO0007311	0.5105	0.0182	1.15	1.18	0.27	C & D
MO0007593	0.3623	0.0260	1.03	1.04	0.37	D
MO0007596	0.4811	0.0252	0.86	0.81	0.55	C

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007600	0.8042	0.0269	1.00	1.06	0.39	D
MO0007625	-0.3771	0.0261	1.00	0.97	0.40	D
MO0007642	-0.8638	0.0191	0.94	0.89	0.43	C & D
MO0007652	0.4983	0.0262	1.09	1.11	0.32	D
MO0007659	-0.9572	0.0271	0.98	0.96	0.39	C
MO0007666	0.2555	0.0250	1.15	1.20	0.27	C
MO0007678	-1.2315	0.0203	0.82	0.70	0.52	C & D
MO0007730	-0.0760	0.0249	0.96	0.95	0.44	C
MO0007741	-0.0669	0.0258	0.87	0.84	0.52	D
MO0007746	0.6572	0.0265	1.13	1.18	0.28	D
MO0007754	0.0606	0.0249	0.98	0.97	0.43	C
MO0007755	-0.5996	0.0185	0.97	0.95	0.42	C & D
MO0007760	-0.9239	0.0193	0.96	0.93	0.40	C & D
MO0007785	0.4253	0.0261	1.05	1.08	0.35	D
MO0007791	0.4192	0.0260	0.97	0.95	0.44	D
MO0007793	0.6794	0.0256	0.88	0.84	0.52	C
MO0007823	-0.9938	0.0195	0.88	0.81	0.47	C & D
MO0007826	-0.3121	0.0252	1.16	1.23	0.24	C
MO0007831	-0.6979	0.0269	0.92	0.87	0.46	D
MO0007841	1.1587	0.0271	1.03	1.13	0.36	C
MO0007860	-0.3574	0.0261	0.99	0.98	0.41	D
MO0007865	0.7094	0.0267	0.79	0.73	0.60	D
MO0007868	0.1350	0.0258	0.93	0.90	0.48	D
MO0007878	1.4704	0.0300	0.89	0.79	0.49	D
MO0007893	0.5929	0.0183	1.20	1.27	0.22	C & D
MO0007939	0.1023	0.0179	0.94	0.93	0.47	C & D
MO0007998	-0.0167	0.0179	0.94	0.92	0.46	C & D
MO0008003	-0.7876	0.0271	0.93	0.87	0.45	D
MO0008143	0.1870	0.0179	0.84	0.80	0.56	C & D
MO0008216	0.8456	0.0260	0.85	0.78	0.56	C
MO0008261	0.1787	0.0252	1.05	1.08	0.36	C
MO0008305	0.5755	0.0263	0.93	0.93	0.47	D
MO0008348	-0.5341	0.0184	1.08	1.14	0.30	C & D
MO0008380	0.1529	0.0258	1.25	1.31	0.17	D
MO0020031	0.0495	0.0249	0.98	0.97	0.43	C
MO0020109	0.2786	0.0250	0.80	0.75	0.60	C
MOA216224	-3.2212	0.0383	0.97	1.12	0.19	C & D
MOA216353	1.2316	0.0274	0.86	0.86	0.51	C
MOA216426	0.7991	0.0259	1.13	1.21	0.28	C
MOA216492	1.6945	0.0315	0.86	0.71	0.50	D
MOA216496	0.8226	0.0259	1.29	1.50	0.12	C
MOA216518	0.4209	0.0251	1.22	1.29	0.20	C
MOA21678	0.2649	0.0180	1.10	1.13	0.32	C & D

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MOA21691	-1.7562	0.0321	0.99	0.98	0.31	C

Table L.5. IRT Item Statistics–Geometry–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007136	1.0987	0.0638	0.99	1.13	0.36	D
MO0007142	0.2228	0.0398	1.06	1.08	0.33	C & D
MO0007197	-0.9046	0.0408	1.01	1.01	0.34	C & D
MO0007227	0.8758	0.0614	0.96	0.97	0.41	D
MO0007248	0.2261	0.0556	0.84	0.80	0.56	C
MO0007598	0.0247	0.0549	1.18	1.27	0.22	C
MO0007645	0.0722	0.0394	1.15	1.19	0.24	C & D
MO0007695	-0.8808	0.0588	0.95	0.91	0.40	D
MO0007708	0.4699	0.0407	0.99	0.99	0.41	C & D
MO0007715	0.8645	0.0613	0.91	0.91	0.47	D
MO0007716	-1.3432	0.0633	0.96	0.92	0.35	D
MO0007736	0.4562	0.0568	1.10	1.14	0.30	C
MO0007745	0.4304	0.0405	1.18	1.24	0.21	C & D
MO0007757	-0.1440	0.0392	0.82	0.77	0.57	C & D
MO0007784	0.1519	0.0396	1.13	1.18	0.25	C & D
MO0007913	-0.6483	0.0574	0.86	0.83	0.50	D
MO0007937	-0.1574	0.0563	1.01	1.00	0.36	D
MO0007946	-1.4070	0.0609	0.88	0.78	0.45	C
MO0007969	-0.0415	0.0548	0.94	0.92	0.46	C
MO0007975	0.6221	0.0592	1.06	1.06	0.32	D
MO0007994	-0.1225	0.0392	0.95	0.93	0.44	C & D
MO0008010	-1.0906	0.0605	0.93	0.88	0.40	D
MO0008029	1.4325	0.0673	0.78	0.61	0.60	C
MO0008046	2.2278	0.0855	0.87	0.65	0.48	D
MO0008060	-0.4522	0.0394	0.96	0.93	0.42	C & D
MO0008072	1.0497	0.0621	1.14	1.21	0.26	C
MO0008082	1.3426	0.0671	1.08	1.21	0.26	D
MO0008098	-0.6212	0.0398	0.91	0.87	0.46	C & D
MO0008180	0.0617	0.0565	1.01	1.05	0.36	D
MO0008206	-0.5893	0.0571	1.25	1.41	0.07	D
MO0008291	0.5286	0.0585	0.93	0.93	0.45	D
MO0008311	-0.1302	0.0392	1.22	1.29	0.16	C & D
MO0008328	0.3573	0.0562	0.98	0.99	0.42	C
MO0008427	0.6985	0.0585	1.01	1.03	0.39	C
MO0008754	-0.6971	0.0555	0.90	0.87	0.47	C
MO0008784	0.4821	0.0569	0.95	0.95	0.45	C
MO0008861	2.2351	0.0857	1.00	0.90	0.34	D
MO0008865	-0.6327	0.0582	0.92	0.89	0.43	D
MO0008871	0.8807	0.0430	0.86	0.83	0.53	C & D

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0008938	1.9949	0.0781	0.97	1.23	0.34	C
MO0008939	0.2482	0.0399	1.14	1.19	0.25	C & D
MO0008946	-0.4719	0.0576	0.90	0.89	0.46	D
MO0008979	-1.9848	0.0508	0.91	0.82	0.37	C & D
MO0008981	1.0197	0.0638	1.00	1.03	0.37	D
MO0009056	0.9833	0.0634	1.03	1.09	0.33	D
MO0009074	-0.2525	0.0563	1.05	1.07	0.31	D
MO0009079	0.2017	0.0578	0.89	0.86	0.50	D
MOG16142	0.8696	0.0429	0.98	0.99	0.40	C & D
MOG16153	-1.0012	0.0572	0.95	0.91	0.41	C
MOG16226	1.1870	0.0455	1.17	1.29	0.20	C & D
MOG16358	1.2906	0.0465	1.00	1.01	0.38	C & D
MOG16431	-0.1673	0.0546	1.19	1.22	0.22	C
MOG16439	-1.4633	0.0616	0.97	0.97	0.35	C
MOG16775	-0.0179	0.0564	1.17	1.20	0.20	D
MOG16788	-0.0205	0.0548	0.83	0.80	0.56	C
MOG16805	0.0368	0.0550	1.12	1.15	0.28	C

Table L.6. IRT Item Statistics–Biology–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007089	0.4261	0.0121	0.97	0.97	0.47	A
MO0007103	-0.2391	0.0144	0.88	0.80	0.52	B
MO0007106	-0.6411	0.0153	0.89	0.86	0.48	B
MO0007109	1.0920	0.0124	0.91	0.85	0.52	A
MO0007112	1.3257	0.0127	1.06	1.06	0.39	A
MO0007117	-0.2320	0.0125	1.15	1.18	0.31	A
MO0007126	0.5907	0.0121	0.94	0.94	0.49	A
MO0007148	0.6049	0.0091	1.06	1.09	0.39	A & B
MO0007157	0.3408	0.0137	0.89	0.86	0.53	B
MO0007189	-0.1307	0.0124	0.88	0.82	0.53	A
MO0007217	-0.0317	0.0141	0.91	0.87	0.50	B
MO0007612	0.6957	0.0136	1.04	1.05	0.41	B
MO0007660	-0.8844	0.0160	0.92	0.87	0.44	B
MO0007675	-0.1296	0.0124	0.89	0.84	0.52	A
MO0007709	1.3361	0.0127	1.26	1.51	0.20	A
MO0007744	1.1252	0.0138	1.14	1.29	0.30	B
MO0007749	1.3143	0.0094	1.26	1.51	0.19	A & B
MO0007763	1.3587	0.0095	1.16	1.31	0.28	A & B
MO0007873	0.1163	0.0122	0.99	0.99	0.45	A
MO0007935	0.2889	0.0138	0.93	0.90	0.50	B
MO0007948	-1.2282	0.0173	0.93	0.86	0.41	B
MO0007965	-0.2416	0.0144	0.91	0.82	0.50	B
MO0007968	-0.5661	0.0130	0.97	0.96	0.43	A

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0008203	-0.1200	0.0093	1.05	1.12	0.38	A & B
MO0008231	-0.5540	0.0099	0.95	0.85	0.46	A & B
MO0008240	-0.9213	0.0105	0.86	0.79	0.49	A & B
MO0008247	0.4904	0.0137	1.05	1.06	0.39	B
MO0008273	0.3294	0.0121	0.94	0.92	0.49	A
MO0008393	-0.2894	0.0095	0.89	0.83	0.51	A & B
MO0008484	0.6370	0.0136	1.02	1.03	0.42	B
MO0008710	0.9942	0.0124	1.02	1.07	0.42	A
MO0008797	0.9073	0.0091	1.02	1.03	0.42	A & B
MO0008840	-0.7914	0.0157	0.90	0.82	0.47	B
MO0008851	0.9423	0.0137	1.26	1.39	0.21	B
MO0008855	-0.8484	0.0159	1.08	1.12	0.31	B
MO0008882	-0.4971	0.0098	0.96	0.87	0.46	A & B
MO0008910	-0.6990	0.0155	0.89	0.81	0.49	B
MO0008915	0.1797	0.0122	0.94	0.92	0.49	A
MO0008957	-0.5532	0.0099	0.85	0.75	0.53	A & B
MO0008963	0.3021	0.0138	1.11	1.16	0.34	B
MO0008972	-0.3045	0.0126	1.00	1.00	0.42	A
MO0008984	0.3539	0.0091	1.25	1.31	0.24	A & B
MO0009018	-0.0541	0.0141	0.95	0.91	0.47	B
MO0009020	-1.0736	0.0142	0.92	0.88	0.44	A
MO0009042	-2.1058	0.0189	0.94	0.80	0.34	A
MO0009119	0.3460	0.0137	0.87	0.83	0.55	B
MO0009202	0.0705	0.0140	1.12	1.16	0.33	B
MO0009214	0.4620	0.0137	1.19	1.24	0.28	B
MO0013608	0.0270	0.0140	0.98	0.96	0.44	B
MO0015259	-0.5512	0.0151	1.05	1.03	0.36	B
MO0015260	1.2743	0.0127	1.24	1.50	0.21	A
MO0015262	-0.6882	0.0154	0.93	0.79	0.47	B
MO0015488	0.3421	0.0121	0.94	0.92	0.50	A
MO0015580	-1.4411	0.0118	0.92	0.81	0.41	A & B
MO0030658	0.1007	0.0092	0.83	0.76	0.58	A & B
MO0030876	-0.6518	0.0158	0.87	0.85	0.50	B
MO0044419	-1.0396	0.0166	0.95	0.93	0.40	B
MO0045927	0.1286	0.0092	0.97	0.95	0.46	A & B

Table L.7. IRT Item Statistics–Physical Science–Dichotomous

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007146	-0.4634	0.0464	0.91	0.88	0.45	A
MO0007170	0.0601	0.0443	0.98	0.98	0.39	A
MO0007179	-0.0345	0.0446	0.92	0.89	0.46	A
MO0007185	0.0380	0.0444	0.92	0.90	0.46	A
MO0007615	0.7658	0.0448	1.08	1.13	0.28	A

Item ID	Rasch Measure	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007626	-0.2333	0.0452	0.96	0.97	0.40	A
MO0007639	0.7106	0.0447	1.03	1.03	0.35	A
MO0007664	-0.3088	0.0456	0.87	0.83	0.50	A
MO0007713	-2.1311	0.0707	0.94	0.93	0.28	A
MO0007734	-0.4134	0.0461	0.94	0.92	0.42	A
MO0007778	-0.3640	0.0458	1.07	1.08	0.28	A
MO0007897	-0.6520	0.0477	0.93	0.89	0.42	A
MO0008284	-0.0042	0.0445	1.12	1.15	0.24	A
MO0008385	0.3960	0.0441	0.84	0.82	0.54	A
MO0008398	-0.2001	0.0451	0.93	0.90	0.44	A
MO0008421	-1.1161	0.0522	0.91	0.83	0.41	A
MO0008483	-0.8675	0.0495	1.01	0.99	0.31	A
MO0008790	0.3980	0.0441	1.02	1.04	0.35	A
MO0008880	-0.2500	0.0453	1.19	1.31	0.14	A
MO0008883	0.4576	0.0442	1.20	1.26	0.16	A
MO0008902	0.2949	0.0441	1.14	1.16	0.23	A
MO0008959	-0.2125	0.0451	0.93	0.93	0.43	A
MO0008987	-0.3768	0.0459	0.94	0.91	0.42	A
MO0009068	0.8255	0.0450	1.03	1.03	0.34	A
MO0009071	0.5213	0.0442	0.97	0.99	0.40	A
MO0009095	0.5613	0.0443	1.15	1.21	0.21	A
MO0009102	-1.5702	0.0588	0.92	0.77	0.38	A
MO0009210	-1.0859	0.0518	0.86	0.73	0.47	A
MO0009224	0.0199	0.0444	0.93	0.92	0.44	A
MO0044943	-0.0123	0.0445	0.99	1.02	0.38	A

Table L.8. IRT Item Statistics–English I–Polytomous

Item ID	Rasch Measure	Threshold 0/1	½	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0001870_S3	-3.1638	-0.5907	0.5907	--	--	0.0516	0.89	0.62	0.38	D
MO0001873_S3	-2.9050	-0.7416	0.7416	--	--	0.0412	0.88	0.63	0.43	C
MO0007207	-0.5153	-1.6624	1.6624	--	--	0.0269	1.18	1.18	0.26	D
MO0007677	0.5527	0.4168	-0.4168	--	--	0.0185	1.50	1.80	0.28	C
MO0008165	0.4663	-0.4705	0.4705	--	--	0.0211	1.08	1.09	0.44	D
MO0008270	0.5014	-2.5764	2.5764	--	--	0.0332	0.97	0.96	0.40	C
MO0008339	-0.2579	-1.8583	1.8583	--	--	0.0265	0.93	0.93	0.50	C
MO0008458	-0.7351	-1.1548	1.1548	--	--	0.0248	1.03	1.03	0.44	D
MO0008726	0.2049	-0.6411	0.6411	--	--	0.0205	1.18	1.20	0.41	C
MO0008779	0.2790	-1.5181	1.5181	--	--	0.0259	1.08	1.09	0.36	D
MO0044056	-0.5393	-1.9676	1.9676	--	--	0.0271	0.88	0.88	0.54	C
MO0044061	0.0198	3.0726	-3.0726	--	--	0.0172	1.05	1.09	0.57	C
MO0001870_S1	0.3312	-3.3097	-0.0797	3.3895	--	0.0261	0.89	0.90	0.53	D
MO0001870_S2	0.2326	-2.6276	-1.1569	3.7846	--	0.0276	0.86	0.84	0.55	D
MO0001873_S1	0.6416	-2.9457	-0.0111	2.9568	--	0.0235	0.85	0.85	0.59	C
MO0001873_S2	0.6101	-2.3335	-1.0688	3.4023	--	0.0228	0.82	0.82	0.61	C

Table L.9. IRT Item Statistics–English II–Polytomous

Item ID	Rasch Measure	Threshold 0/1	½	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0001817_S3	-3.1531	-0.5912	0.5912	--	--	0.0233	0.89	0.58	0.37	D
MO0001843_S3	-3.2870	-0.5789	0.5789	--	--	0.0216	0.88	0.57	0.39	C
MO0007296	0.3269	-0.6461	0.6461	--	--	0.0093	1.16	1.20	0.41	D
MO0007842	0.7111	-0.9143	0.9143	--	--	0.0098	1.37	1.43	0.21	D
MO0008070	-0.8428	0.9991	-0.9991	--	--	0.0095	1.09	1.17	0.47	D
MO0008146	-0.7939	-1.7628	1.7628	--	--	0.0118	0.93	0.93	0.49	D
MO0008162	0.7515	-0.6439	0.6439	--	--	0.0088	1.15	1.16	0.44	C
MO0008185	0.0811	-2.4663	2.4663	--	--	0.0146	0.96	0.95	0.40	D
MO0009103	1.0737	-0.5736	0.5736	--	--	0.0090	1.15	1.16	0.42	C

Item ID	Rasch Measure	Threshold 0/1	½	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0044085	-1.8459	-1.2854	1.2854	--	--	0.0115	0.96	0.99	0.46	C
MO0001817_S1	0.6164	-2.9771	-0.1581	3.1352	--	0.0112	0.87	0.87	0.57	D
MO0001817_S2	0.5483	-2.3928	-1.0365	3.4293	--	0.0110	0.83	0.82	0.60	D
MO0001843_S1	0.3661	-3.2424	0.1954	3.0470	--	0.0105	0.82	0.81	0.63	C
MO0001843_S2	0.2468	-2.4302	-1.1420	3.5722	--	0.0108	0.82	0.81	0.61	C

Table L.10. IRT Item Statistics–Algebra I–Polytomous

Item ID	Rasch Measure	Threshold 0/1	½	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007870	-0.4126	-0.4963	0.4963	--	--	0.0087	0.89	0.88	0.62	C
MO0008411	0.2638	0.1107	-0.1107	--	--	0.0092	1.06	1.06	0.55	D
MO0008746	-0.6209	-1.6908	1.6908	--	--	0.0108	1.02	1.03	0.46	C
MO0008766	0.7351	-0.6392	0.6392	--	--	0.0105	1.33	1.38	0.33	D
MO0008925	0.2370	0.0651	-0.0651	--	--	0.0092	1.26	1.32	0.44	D
MO0008998	0.1058	0.0708	-0.0708	--	--	0.0084	0.82	0.75	0.67	C
MO0009019	-0.3864	0.5737	-0.5737	--	--	0.0085	0.90	0.87	0.62	D
MO0009069	-0.6722	-0.9214	0.9214	--	--	0.0101	0.96	0.95	0.54	D
MO0008731	-0.7257	-0.7892	0.1364	0.6528	--	0.0070	0.90	0.89	0.68	C
MO0008734	-0.1189	-0.6502	0.1224	0.5279	--	0.0069	0.94	0.93	0.68	C
MO0008762	1.3923	0.0529	0.0765	-0.1293	--	0.0101	0.96	0.82	0.60	D
MO0008767	-0.3098	-1.6189	-0.5682	2.1871	--	0.0091	1.13	1.12	0.44	D
MO0008732	0.3802	-0.3195	-0.5157	0.8373	-0.002	0.0060	1.09	1.03	0.69	C

Table L.11. IRT Item Statistics–Algebra II–Polytomous

Item ID	Rasch Measure	Threshold 0/1	1/2	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007275	-1.2145	0.259	-0.259	--	--	0.0192	0.86	0.75	0.53	C
MO0007603	0.6887	1.6528	-1.6528	--	--	0.0157	1.14	1.37	0.52	C
MO0007843	1.1406	-1.2521	1.2521	--	--	0.0147	1.20	1.21	0.34	C & D
MO0007915	0.3501	5.1051	-5.1051	--	--	0.0153	1.43	1.91	0.42	D

Item ID	Rasch Measure	Threshold 0/1	1/2	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0008067	0.9097	-0.8303	0.8303	--	--	0.0190	0.82	0.79	0.64	C
MO0008260	-0.7163	0.7494	-0.7494	--	--	0.0171	0.90	0.79	0.58	D
MO0008372	-0.6295	-0.7204	0.7204	--	--	0.0191	1.10	1.16	0.43	D
MO0007770	1.0923	-0.9806	-0.9638	1.9444	--	0.0158	1.31	1.41	0.43	D
MO0008281	0.8661	-1.8335	1.2920	0.5415	--	0.0177	1.15	1.13	0.47	D
MO0008298	0.9668	-1.3881	-0.6877	2.0758	--	0.0164	1.03	1.02	0.56	D
MO0008314	-0.9353	-0.7869	-0.1327	0.9196	--	0.0155	0.88	0.87	0.62	C
MO0020625	0.6163	0.3049	1.5584	-1.8633	--	0.0126	1.26	1.46	0.59	C
MO0008066	-0.9689	-0.8667	-1.1419	2.2078	-0.1991	0.0129	1.06	1.12	0.61	C

Table L.12. IRT Item Statistics–Geometry–Polytomous

Item ID	Rasch Measure	Threshold 0/1	1/2	2/3	3/4	Standard Error	Infit	Outfit	PTMA	Core Form
MO0007628	0.0229	-0.3225	0.3225	--	--	0.0369	0.93	0.92	0.58	C
MO0007911	-0.5962	-0.7442	0.7442	--	--	0.0400	0.97	0.94	0.52	C
MO0007943	0.3415	-0.1525	0.1525	--	--	0.0381	1.08	1.06	0.49	D
MO0007944	-1.2832	-0.9374	0.9374	--	--	0.0440	0.92	0.86	0.51	C
MO0008312	0.7135	-0.1473	0.1473	--	--	0.0398	0.86	0.90	0.61	C
MO0008320	-0.1949	-0.1582	0.1582	--	--	0.0369	0.95	0.92	0.55	D
MO0008829	1.2736	-1.3632	1.3632	--	--	0.0473	0.97	0.97	0.51	C
MO0009032	-0.5900	1.2661	-1.2661	--	--	0.0345	1.00	1.05	0.50	D
MO0009046	0.7007	-0.0233	0.0233	--	--	0.0400	1.25	1.40	0.38	D
MO0007978	1.4209	-2.6214	0.3595	2.2619	--	0.0444	0.92	0.91	0.56	C
MO0008022	0.7130	0.7698	-0.1539	-0.6159	--	0.0311	1.04	0.96	0.66	C
MO0009125	-0.3171	-0.5518	-0.1061	0.6579	--	0.0306	0.80	0.78	0.67	D

Table L.13. IRT Item Statistics–Biology–Polytomous

Item ID	Rasch Measure	Threshold				Standard Error	Infit	Outfit	PTMA	Core Form
		0/1	1/2	2/3	3/4					
MO0007150	-0.4859	4.6826	-4.6826	--	--	0.0095	0.98	1.08	0.55	B
MO0007160	0.3582	-0.6188	0.6188	--	--	0.0085	1.11	1.14	0.52	A
MO0007177	-0.6925	6.016	-6.016	--	--	0.0084	0.97	1.20	0.53	A
MO0007219	1.2023	-0.0991	0.0991	--	--	0.0092	1.15	1.19	0.51	B
MO0007232	1.3944	0.1997	-0.1997	--	--	0.0085	1.03	1.27	0.55	A
MO0007254	0.4979	0.325	-0.325	--	--	0.0077	0.93	0.95	0.64	A
MO0007305	0.1476	0.4443	-0.4443	--	--	0.0089	1.10	1.10	0.56	B
MO0008904	-0.6968	0.2461	-0.2461	--	--	0.0088	1.03	1.23	0.49	A
MO0008970	0.5762	1.2302	-1.2302	--	--	0.0075	0.91	0.85	0.66	A
MO0009096	-1.5859	0.0222	-0.0222	--	--	0.0116	0.86	0.74	0.49	A
MO0009144	-1.2492	0.1023	-0.1023	--	--	0.0103	0.76	0.56	0.58	A
MO0009226	0.4192	-2.2289	2.2289	--	--	0.0132	0.92	0.90	0.51	B
MO0013657	0.5503	0.2016	-0.2016	--	--	0.0088	1.05	1.05	0.58	B
MO0014514	-1.4309	-1.8471	1.8471	--	--	0.0112	1.09	1.07	0.41	A

Table L.14. IRT Item Statistics–Physical Science–Polytomous

Item ID	Rasch Measure	Threshold				Standard Error	Infit	Outfit	PTMA	Core Form
		0/1	1/2	2/3	3/4					
MO0007121	2.4024	-1.2226	1.2226	--	--	0.0420	0.88	0.83	0.52	A
MO0007638	0.3315	0.3391	-0.3391	--	--	0.0274	0.96	0.93	0.57	A
MO0007698	0.5753	-0.7154	0.7154	--	--	0.0313	1.11	1.12	0.42	A
MO0007728	0.4858	0.0060	-0.0060	--	--	0.0284	1.13	1.17	0.45	A
MO0007804	0.4830	-0.7728	0.7728	--	--	0.0315	1.19	1.22	0.35	A
MO0007820	0.6164	4.3930	-4.3930	--	--	0.0256	1.05	1.06	0.56	A
MO0008106	0.7941	-1.6264	1.6264	--	--	0.0380	1.00	1.00	0.43	A
MO0008212	-1.1761	-1.0672	1.0672	--	--	0.0378	0.94	0.98	0.45	A
MO0008299	0.0365	-0.1403	0.1403	--	--	0.0289	0.99	1.02	0.52	A
MO0022372	0.6983	-0.6920	0.6920	--	--	0.0314	0.99	0.99	0.51	A